Analyses of microbial communities at biogeochemical interfaces and their interaction with organic pollutants in soil

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig
zur Erlangung des Grades eines Doktors der Naturwissenschaften
(Dr. rer. nat)
genehmigte
Dissertation

Von Guo-chun Ding
aus Jiangsu, China

1.Referentin oder Referent: Prof. Dr. Kornelia Smalla
2. Referentin oder Referent: Prof. Dr. Dieter Jahn
eingereicht am:12. 07.2011
mündliche Prüfung (Disputation) am: 05.12. 2011
Druckjahr 2012

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

Ding GC, Heuer H, Zuhlke S, Spiteller M, Pronk GJ, Kogel-Knabner I and Smalla K. (2010) Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes by using a novel PCR detection system. Appl Environ Microbiol 76: 4765-4771.

Weinert N, Piceno Y**, Ding GC**, Meincke R, Heuer H, Berg G, Schloter M, Andersen G and Smalla K. (2011) PhyloChip hybridization uncovered an enormous bacterial diversity in the rhizosphere of different potato cultivars: many common and few cultivar-dependent taxa. FEMS Microbiol Ecol 75: 497-506.

Poster

Guo-Chun Ding, Holger Heuer, Sebastian Zühlke, Michael Spiteller, Geertje Johanna Pronk, Katja Heister, Ingrid Kögel-Knabner, Jianping Xie, Jizhong Zhou and Kornelia Smalla. Responses of microbes in two typical central European soils to phenanthrene spiking as revealed by PCR targeting PAH-RHDα genes, GeoChip, PCR-DGGE and amplicon pyrosequencing(2011). BAGECO 11, Greece, 29/05-2/06

Guo-Chun Ding, Holger Heuer and Kornelia Smalla. Similar ribotypes belonging to *Sphingomonadales* and *Burkholderiales* enriched in two different phenanthrene spiked soils as revealed by pyrosequencing and DGGE (2010). Soil metagenomic, Germany, Braunschweig, 08/12-10/12

Guo-Chun Ding, Holger Heuer and Kornelia Smalla. Diversity and abundance of PAH-ring hydroxylating dioxygenase (*pah-rhdα*) genes in two phenanthrene contaminated soils differed as revealed by a novel PCR detection system (2009). Biological responses to chemical contaminants: from molecular to community level. Portugal, Aveiro, 2/09-4/09.

Guo-chun Ding, Holger Heuer, Sebastian Zühlke and Kornelia Smalla. Analyses of microbial communities at biogeochemical interfaces and their interaction with organic pollutant in soil (2009). BAGECO 10, Sweden, Uppsala, 29/05-2/06

Summary

To study the responses of soil indigenous microbial communities to persistent organic pollutants, the Eutric Cambisol (60°N, 17°E) and the Luvisol (48°N, 11°E) were spiked with the model polycyclic aromatic hydrocarbons (PAH) substrate phenanthrene and incubated at room temperature. Total community DNA was extracted from samples collected at Day 0, 21 and 63 and used for further molecular analysis.

The key step of microbial PAHs degradation is catalyzed by dioxygenases which incorporate molecular oxygen into the aromatic nucleus. Bacteria harboring such genes could play an essential role for microbial PAHs mineralization in soils. Due to narrow detection spectrum or short amplicon sizes of previously published PCR systems, a novel PCR based on all 40 different *PAH*-ring hydroxylating dioxygenases (*RHD)α* sequences deposited in GenBank (as of December 2007) was developed and confirmed experimentally and by *in silico* analysis. This system was applied to study the responses of bacteria carrying *PAH-RHDα* genes to phenanthrene in both soils. *PAH-RHDα* amplicons could be detected from the Luvisol at Day 21 but not from the Cambisol. Quantitative real-time PCR based on the same primers revealed that 63 days after phenanthrene spiking, *PAH-RHDα* genes were one order of magnitude more abundant in the Luvisol than in the Cambisol, while they were not detected in both control soils. Cloning and sequencing showed the dominance of *phnAc* genes in the contaminated Luvisol. In contrast, high diversity of *PAH-RHDα* genes of Gram-positive and Gram-negative bacteria was observed in the phenanthrene-spiked Cambisol. In conclusion, sequence analysis of the amplicons obtained confirmed the specificity of the novel primer system and revealed a soil type-dependent response (abundance and diversity) of *PAH-RHDα* gene-carrying soil bacteria to phenanthrene spiking.

However, just to understand the key genes for PAHs mineralization is not enough to evaluate the influences of organic pollutants on other soil functions. GeoChip developed by He et al consists of 24243 oligonucleotide probes targeting more more than 10,000 genes in more than 150 functional groups allows stimutaneously studying nitrogen, carbon, sulfur and phosphorus cycling, metal reduction and resistance, and organic contaminant degradation (He et al 2007). Samples collected

at Day 0 and 21 were analyzed by GeoChip. More genes were detected from contaminated soils which could due to increased relative abundance of bacteria in the polluted lines. Five hundred and nineteen genes were identified with significantly different signal intensity between contaminated Luvisol and its corresponding control and these genes could link to distant taxonomic groups as well as different function groups. Among these discriminative genes, most organic remediation genes links to the degradation of one ring aromatic hydrocarbons, but six genes (*nahAa,nahAc, nahQ* and *narR*-like) were identified as upper pathway genes for PAHs degradation. In summary, GeoChip analysis showed strong responses of microbial community to phenanthrene spiking in Luvisol.

IncP-9 plasmids are known to carry operons which encode the upper and lower pathway for naphthalene degradation. These plasmids are assumed to contribute to the adaptation of *Pseudomonas* species to pollution. A novel primer system targeting conserved stretches in the *rep-oriV* region was developed based on the sequences of 28 different IncP-9 plasmids belonging to eight groups. The specificity of the primer was confirmed experimentally as no amplicons were acquired against IncP-1, IncN, IncW and IncQ. This system was also combined with Southern blot to increase the specificity and sensitivity for detecting incP-9 plasmids from total community DNA from soil. Regardless of phenanthrene contamination, amplicons with expected size were acquired for all samples at Day 63. Interestingly, three amplicons did not hybridize with pooled probes generated from typical incP-9 plasmids (pM3, pBS265, pMG18, pNL60, pNL15, pSVS15, pNF142, pSN11, P80 and pWWO). This result suggests that novel incP-9 plasmids which might not involve in PAH degradation were present in our soil samples.

GeoChip analysis showed the majority of ORGs with significant increased signal intensity are involved in one ring aromatic hydrocarbons degradation. This finding suggests that microbes with lower pathway may be major part of enriched bacteria after phenanthrene spiking. Partially to answer this hypothesis, 16S RNA gene amplicons were quantified by real time PCR and the community structure was analyzed by PCR-DGGE and pyrosequencing. The analysis based on community profiles showed that soil type, bacterial taxonomic group and exposure time dependent responses to phenanthrene contamination. The relative abundance of

Sphingomonadales and *Polaromonas* were significantly increased in both polluted soils. Decreased richness of detected OTUs was also observed for both phenanthrene-spiked soils. Strikingly, 30 operational taxonomic units (OTUs) were enriched in both contaminated soils which originate from distant sites. Microbes with the lower microbial PAH degradation pathway could be the major enriched bacteria because the relative abundance of enriched taxa was much higher than the relative abundance of PAH-RHDα genes in both soils. In conclusion, soil type, exposure time and taxonomic group dependent responses to phenanthrene-spiking were observed for all studied bacterial communities by DGGE analysis of 16S rRNA gene amplicons, and barcoded 16S rRNA gene pyrosequencing revealed the enrichment of ribotypes belonging to *Sphingomonadales* in both polluted soils.

So far, the responses of indigeous microbial community in both model soils to phenanthrene were analyzed in a comprehensive manner. These studies provided essential information and tool sets (primers targeting PAH-RHDα genes and oriV-rep regions of IncP-9 plasmids and pipelines for processing data generated by GeoChip and barcoded pyrosequencing for 16S rRNA genes) for the joint experiment which targets the interaction of soil microbial communities with persistent organic pollutants at biogeochemical interfaces in soil. But the information on biogeochemical interfaces in soil is rarely known. Attempts were made to simulate them with artificial soils of different minerals composition. To explore the influences of mineral composition on the microbial community assembled, batch experiment was performed with seven artificial soils inoculated with the same soil microbial extracting and autoclaved manure. Samples were taken 1, 9, 31 and 90 days after inoculation. Bacterial and fungal communities were monitored by PCR-DGGE analysis of partial 16S rRNA gene and ITS region respectively. Charcoal was found to strongly influence the reassembly of bacterial communities. Effects of montmorillonite, illite and iron oxides on microbial communities could clearly be observed at day 90. Thus, four artificial soils (QMC, QM, QI, and QIF) were selected for further intensive studies. Real time PCR showed that there were no significant differences in 16S rRNA gene copy number was detected between any pair of sampling times and artificial soils. Pyrosequencing analysis for the four artificial soils samples collected at Day 90 confirmed the strongest influence of charcoal on the bacterial community which has a

low detected diversity. *Sphingomonas*, *Lysinibacillus* and *Streptomyces* were significantly enriched in QMC.

Reference

He, Z., T. J. Gentry, C. W. Schadt, L. Wu, J. Liebich, S. C. Chong, Z. Huang, W. Wu, B. Gu, P. Jardine, C. Criddle, and J. Zhou. 2007. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J **1:**67-77.

Introduction to the PhD thesis

This PhD thesis is a part of a cooperative project named “Biogeochemical interfaces in soil” funded by Deutsche Forschungsgemeinschaft (DFG) to systematically study the structure and function of biogeochemical interfaces in soil and to unravel their effect on the fate of organic chemicals.

Firstly, the responses of soil microbial communities to model PAH substrate phenanthrene were studied. In chapter 3, a novel PCR system which targets wide spectrum PAH-RHDα genes was developed and applied to study the responses of bacteria carrying *PAH-RHDα* to phenanthrene spiking (Ding et al., 2010). Chapter 4 summarizes the insights gained from GeoChip analysis of soil samples collected at day 0 and 21 after contamination in order to detect changes of functional genes. To explore the responses of bacterial population carrying IncP 9 plasmids in both soils, a PCR system targeting at *oriV-rep* regions in the backbone was described in chapter 5. The responses of bacterial and fungal community to phenanthrene spiking were explored by PCR-DGGE and pyrosequencing analysis based on 16S RNA gene amplicons in chapter 6.

Secondly, the influence of mineral composition and charcoal on the reassembling of bacterial community in artificial soils was studied by the PCR-DGGE and pyrosequencing of 16S rRNA gene amplicons. This study pointed out importants elements which might influence the formation of biogeochemical interface in soil.

Thirdly, Chapter 1 is a general introduction to background, major questions and popular molecular methods in soil microbial ecology. To keep the flow of rational, some parts (cited as chapter 2) were based on the book chapter (chapter 2) by Kornelia Smalla, Holger Heuer and the author of this thesis.Chapter 2 is a book chapter (published) which provides an overview of cultivation-independent detection of genes present in soil bacteria. Finally, a general outlook was made. In supplement chapter (published) is an example for efforts invested on analyzing the high dimentional data genereated by a phylogenetic microarray termed “Phylochip“.

Chapter 1: General introduction

1.1 Soil is a complex platform for microbes

Soil is a very complex natural matrix comprised of solids (minerals and organic matter), liquid, and gases. Quartz, clay, oxides (Fe, Mn, Al) and carbonates are the major solid minerals. These minerals are glued together into irregular porous microaggregates by organic matter derived from plants, animals or microbes. The surface area can vary from 11 cm^2 up to 8 million cm^2 per gram of soil [1] due to different proportions of sand, silt, and clay. Highly heterogeneous distribution of water, nutrient and pores forms diverse biogeochemical interfaces which affect the composition and activity of microbial community. Temporal binding of microaggregates into macroaggregates by plant or hyphae further increases the complexity of soil structure. Beside the binding effects, plant roots also exude a large amount of organic substances which serve as food for a variety of microbes. The soil influenced by the roots of a plant is termed ‘rhizosphere’. The activity and density of bacteria in the rhizosphere is much higher than in the surrounding bulk soils [2,3,4,5,6]. Soil animals such as earthworms also alter the chemical and physical characters of soils by digestion and drilling. They convert organic matter into stable humus colloids which they transport to lower soil horizons. The borrowing process increases soil aeration, drainage and porosity. These activities create or change the biogeochemical interfaces which could strongly influence the community structure and activities of microbes [7,8,9]. Except for biological influences, dynamic processes between atmospheres, hydrosphere also constantly influence the quality and quantity of these interfaces.

The diversity of soil bacteria is huge. An extreme variety of dynamical biogeochemical interfaces are believed to be the major driving forces for this diversification. Bacteria with best fitting capacity may occupy, and propagate at, certain niches. Examples for “the environment selects” have been shown for several niches. For instance, enrichments of bacteria in rhizosphere were found dependent on the plant species, cultivars, developing stages of plants as well as soil types [2,10,11]. Increased abundance of PAH degraders were frequently found in the oil, coal tar contaminated, or other PAH contaminated soils [12,13,14,15]. Several factors such as types of PAHs, exposure time, salinity, and temperature could also affect the

abundance and diversity of enriched bacteria [12,16,17,18,19]. The complex porous structure not only provides varied biogeochemical interfaces but also supplies hiding niches. Bacteria mainly colonize pores with sizes of 0.8-3 μm where they are protected against grazing while macropores contain fewer bacteria. In addition, the activities of soil animals, interactions between atmospheres, hydrosphere and so on, not only create new niches, but also transport microbial communities from one niche to another. These actions could also contribute to the tremendous diversity of soil bacteria. Another contribution to the huge diversity of bacteria in soils could be the oligotrophic conditions which do not favor the dominance of few bacteria.

The functions of soil bacteria are diverse and vital in the flow and conservation of energy and material in terrestrial ecosystems. The bacteria involved provide many ecosystems with great flexibility, stability and efficiency. For example, when a legume plant is highly demanding for nitrogen during its growth, species of *Rhizobia* could fix nitrogen from the atmosphere indirectly using the energy fixed by legume plants. In addition to supplying the platforms for exchanging of materials, bacteria in the rhizosphere could increase the competence of plants in several other aspects such as enhancing their ability to defend themselves against pathogens. A huge amount of organic material with varied structure is produced and released into soils. Bacteria, as well as fungi and other decomposers, convert these organic materials into inorganic forms which can be taken up by plants. Soil bacteria play many ecological roles, and many of their functions are of high redundancy. The levels of redundancy might relate with the frequency of exposures to certain niches. For example, bacteria with the ability to degrade low molecular weight aromatic hydrocarbons such as benzene, ethylbenzene, xylene and toluene (BEXT) are prevalent in unpolluted samples [16]. But bacteria degrading high molecular weight were frequently found at heavily contaminated sites [24, 81].

In summary, heterogeneous physicochemical structure of soils, together with several dynamic influences from plants, animals or surroundings, support a huge diversity of microbes, which are of several essential and/or unnecessary ecological functions**.** The giant project aims at the characterization of these biogeochemical interfaces in soil, and ultimately attempts to draw dynamic interlinks with joint interdispline power.

1.2 Questions to answer in soil microbial ecology

Soils are labyrinths with huge amounts of different microbes whose diversity and functions only have partially been uncovered. So far it is believed that heterogeneous distribution of nutrients and niches are a major force driving the microbial diversity. But the detailed mechanisms are still hidden. Several studies tried to link soil parameters such as pH and salinity with the bacterial diversity [20,21,22,23]. But such linking is still too weak due to the fact that only a very small fraction of bacteria was retrieved and studied. Solid linking could only be possible when the diversity of soil bacteria could be reliably detected or estimated. Therefore, unraveling the diversity of soil bacteria is still a big challenge in microbial ecology.

Beside the diversity of soil bacteria, the total functional potential of certain soils is hardly known. Genomic studies provide a chance to understand what a bacterium can do. Nevertheless most studies still focus on cultivable bacteria [24,25,26]. But it is known that in soil the majority of bacteria are viable but uncultivable with traditional methods [27]. Genome analysis at single cell level could help overcome such obstacles [28,29,30]. No question that this work will be extremely difficult due to the huge diversity of soil bacteria, many of which hide in niches preventing isolation. Direct studies on soil total community DNA could help estimate the total functional potential, even though the boundary of each functional unit could be blurred.

Every living microbe is an open system which interacts with its surroundings. Therefore, the soil microbial community as a total is also an open system which not only shifts along time but also responds to the environmental changes. The influence of plants, animals, physiochemical change of structure and function of soil microbial communities attracted a lot of attention [2,7,9,20,23]. Qualitative or semi-quantitative linking between microbial structure or few functional groups and environmental parameters were made. But the black box for the complex reaction of microbes at different dynamic environmental conditions and the extremely complicated interactions among microbes is still almost closed. Micro-scale level functional and structural studies on microbial communities as well as corresponding real-time physicochemical measurements are needed to uncover such relationship.

Knowledge on vertical and horizontal evolution of soil microbes is also important for understanding their versatility, sensitivity and diversity [31,32]. The fast evolution rate or spreading genes among different bacteria allow them to adapt to the changing environmental conditions quickly. Changes in microbial communities are not only limited in quantities of microbes but also in diversity. Mechanisms related with quality changes of microbes are another focus of microbial ecology.

In summary, the diversity, structure, functional potential and evolution of soil microbial communities as well as the dynamic linking with surroundings are great interests to soil microbial ecologists.

1.3 Nucleic acid based methods to study soil microbial communities

1.3.1 Sampling and nucleic acid isolation (chapter 2)

Most studies were performed in reductive fashion because of applicability. But heterogeneity in soils cannot be simply ignored. Sampling strategies mainly depend on the hypothesis tested, but also should reflect the variability of the studied ecosystems. Pre-experiments could help assist in determining the numbers of samples to be analyzed. Frequently, composite samples were used to represent certain ecosystems. This kind of representative manner makes the studies applicable, but could also result in an underestimation of the meaningful biological variance within systems.

A prerequisite for studying microbial communities by molecular methods is extracting nucleic acids from the soil matrix. Even though protocols for soil DNA extraction have been optimized during the last thirty years, none of them is suitable for all kinds of soils, especially for those under heavy contamination. Criteria for evaluating different protocols are high molecular weight and purity. Due to the complex structure of niches in soil and varying abilities of different microbes to resist cell lysis, efficient lysis of the bacterial, archaeal and fungal cell walls is a crucial step for recovering soil representative DNA. But the strength of lysis also needs to be a trade-off as too rigorous lysing methods might shear DNA released from cells that are easy to lyse. Sheared DNA can cause PCR artefacts and is not suitable for direct

cloning of large DNA fragments. Depending on the purposes, the strength of lysis needs to be adjusted. The complete removal of co-extracted humic acids is critical for the subsequent molecular analysis as humic acids were shown to interfere with DNA hybridization, restriction enzyme digestions and PCR amplification [33].

Commercial kits for total DNA extraction from soil facilitated largely the procedure in particular in view of a simplification and miniaturization of the method. Soil DNA extraction kits are often less time consuming and more efficient in removing co-extracted humic acids. The DNA yield might vary considerably for different DNA extraction kits used for the same soils. A recently performed comparison of two frequently used soil DNA extraction kits revealed that the 16S rRNA gene copy numbers were by approximately two orders of magnitude different (Chapter 2). Differences in DNA yield might be no problem if only dominant community structures were compared. But it could cause very serious problems when the community structures of less abundant taxa or genes were targeted. Therefore, kits of high recovering efficiency are recommended.

1.3.2 PCR and quantitative Real-Time PCR (Q-PCR)

Polymerase chain reaction (PCR) is a fundamental method in modern molecular biology and microbial ecology. Many studies are based on PCR amplicons. Therefore, the merits, limitations and related influenced factors of this technique must be understood. The spectrum of the PCR is mainly determined by primers. Many previous primers were designed based on a limited number of sequences which do not cover enough diversity of certain genes. Therefore, the spectrum of the PCR amplification system should be checked before application. Updating of the amplification system is needed if it is necessary. Amplicons for 16S rRNA gene were frequently used to study the diversity and structure of bacterial and archaeal communities in different environmental samples. Several universal primers targeting this gene have been published. Recently, Wang and Qian checked the coverage of previously published primers and found that some of them are of very low coverage. They also suggested conserved regions such as A519, E969-983, E1063-1081, U515 and E517 for metagenomic studies [34]. To widen the scope of PCR systems, degenerated primers were frequently used. But a high degree of degeneracy could

also lead to low amplification efficiency. Since the last 5 or 6 bp of 3' end of primer is critical for the initiation of the reaction, reduced degeneracy could be reached by allowing mismatch at the 5' end. Meanwhile, this strategy could also be used to prevent hairpin or dimer structures of the primer. But distortion of community structure by degenerated primers should also be noticed. Compared to those gene types with low relative abundance, the amplification efficiency for dominant gene types could decrease earlier due to the exhaustion of primers [12].

So far, there is no good way to estimate the efficiency of soil DNA extraction. Therefore, precise quantification of the copy number of certain genes in the soil sample is still an impossible work. But in general, equal or comparative DNA extraction efficiency was expected if the same protocol was used. Hence, imprecise quantification could still give ecological indications. To quantify genes in certain DNA samples, a serial dilution of temperate DNA solutions with known copy number were used as reference. The expected copy numbers in unknown samples should be within the range of the serial dilution of temperate DNA. The assumption of this estimation is that equal amplification efficiency was expected between standard gene and samples. Several parameters, such as quality of DNA samples, or mismatch between target genes and primers, could influence the efficiency of PCR. Both reproducibility and efficiency of the samples could be estimated by serial dilutions of the sample DNA. Theoretically, the copy number in the samples could be more precisely estimated by this way. To measure the amount of amplicons in the samples, a fluorescent signal which is proportional to the amount of amplified products is needed. The most used fluorescent dye is SYBR green 1 which could bind the double-strand DNA and emits at 520 nm. The disadvantage of this application is that primer dimer as well as unspecific amplification could introduce bias in quantification. Therefore, on the one side, the primers need to be carefully chosen; on the other side, the target genes in the samples should be of high abundance which prevents unspecific amplification. Of course, the melting curves should be used to check the specificity of the amplification. Specific quantification could be reached by using Taqman probes of which attach fluorophores at the 5' end and a quencher to the 3' end. During the annealing step, the probe binds to the templates. Then Taq polymerase separates the fluorophores and quencher during extension. Therefore the fluorescence signal of fluorophores could be detected. But

the huge and unknown diversity of target genes prevents this kind of specific detection frequently.

1.3.3 Methods to study microbial community structure

1.3.3.1 Fingerprinting techniques

Fingerprinting techniques such as DGGE, TGGE, and SSCP are intensively applied to compare the microbial communities in different treatments. The merits of these methods are that they are relative cheap, reproducible and suitable for large numbers of samples. In general, the structure of microbial community could still be reflected on the fingerprints. Therefore, by analyzing the profiles, it is possible to compare the community structure. Permutation tests using the similarity matrix could further confirm the significant differences between treatments in microbial community structure when there are enough replicates (>3). Taxonomic information for interested bands could be acquired by cloning and sequencing the amplicons from the excised bands. But the rate to successfully identify the distinguished bands somewhat depends on luck. Also, detailed taxonomic information on studied microbial community structure is hardly known. From the technical aspects, a huge amount of details are needed to be considered to generate profiles of high quality. For example, heterogeneous double-strand DNA, which forms at the plateau of PCR due to the exhaustion of primers, should be minimized. The amounts of PCR product loaded in the DGGE gel need to be well estimated, as too much will lead to heavy background; too less is not enough to present the community structure. Staining the gel and further analysis of profiles are also heavily dependent on experience. In summary, fingerprinting techniques such as DGGE are suitable for studying complex microbial community but a lot of experience is needed. We applied this convient, mature and affordable technique to monitor the community shift after phenanthrene spiking (chapter 6) and to make primary comparison of microbial community assembled in different artificial soils (chapter 7).

1.3.3.2 PhyloChip analysis

Recently, 16S rRNA gene fragments amplified from soil DNA were also analyzed by so-called PhyloChips that enable the parallel detection of more than 8,700 operational taxonomic units (OTU) and the ability to identify individual OTU varying

by five orders of magnitude in abundance [35,36]. Typically more than 2,000 OTU can be simultaneosly be detected in soil DNA. Due to the amount of data acquired, data mining was demanding. In the study by DeAngelis *et al.* (2009) the rhizosphere effect was studied by PhyloChip analysis [37]. OTU responding to the root exudates were detected in 17 of the 44 phyla. PhyloChip analysis of 16S rRNA gene fragments enables a comparison of multiple samples and a determination of taxa responding to biotic or abiotic factors. The great advantage of PhyloChips compared to other fingerprinting methods is that information on the taxonomic affiliation of these responders can be obtained (chapter 2). But some technical aspects need to be noticed. First of all, amplicons with different annealing temperatures from 48°C to 58°C were pooled for further analysis. This procedure is expected to be a trade between increasing diversity and loss of relative abundance. Therefore, horizontal comparison of OTU signals between different treatments was suggested. Secondly, only those OTU present on the PhyloChip could be detected. Thirdly, further analysis such as Q-PCR on responded taxa is needed to confirm the findings. In summary, PhyloChip is a very powerful tool to analyze complex soil microbial communities, and that provides insights into the type of bacteria present but not on their relative abundance.

1.3.3.3 Pyrosequencing

The fast developing sequences techniques make it possible to generate a huge amount of sequences at affordable prices. More than one million high-quality sequences with average length of 400 bp could be generated in one run of ca 10 hours by GS FLX Titanium sequencer. Pyrosequencing of 16S rRNA gene amplicons became a very popular method to study the soil microbial community [38]. Compared with previous indirect way to study 16S rRNA amplicons, this method supplies more reliable taxonomic information on the total community as well as the responders. One challenge for mining biological meaning from the huge amount of sequences was multiple alignments. Traditional tools such as clustalW were not suitable for this job as the time costs increase in quadratic models while the increase of computer power was much slower. Pioneers have developed powerful tools such as RDP multiple aligner, Mothur, which increased the efficiency of the computation very significantly [39,40,41]. In addition to applying new algorithms, parallel programming strategies also accelerated calculation by efficiently using modern computers with multiple processors and multiple computers. Then assigning sequences into OTU at different

DNA distances is a popular method. DNA distance of 16S RNA genes reflected different taxonomic levels. In general, sequences of less than 3% DNA distance are typically assigned to the same species, those with less than 5% distance are typically assigned to the same genus, and those with 20% DNA distance are typically assigned to the same phylum [42]. BlastN from NCBI and Naive Bayes identifier from RDP were frequently used to assign the taxonomic information to the sequences. The first tools try to find the closest sequences in the Genbank. But due to uncertainty or lack of taxonomic information of certain sequences, this method is only used for a limited number of sequences. The latter tool could identify sequences to genus level and give confidential level for each classification. Compared to the BlastN, less computation is needed and therefore it is suitable for datasets with big amounts of data. Compared to clone library and PhyloChips, pyrosequencing provides in-depth and more reliable data on the complex microbial community structure. However, despite the enormous amount of data acquired by pyrosequencing, ten thousand sequences are still a very small fraction of $> 10^9$ bacteria living in one gram of soil. Therefore, more attention should be paid to the statistical exploration of these datasets (chapter 2). First of all, how well does the dataset reflect the real diversity? Table 2.1.2 demonstrates the relationship among the relative abundance of a species, the total number of sequences and the chance to detect this species. For example, the relative abundance of a species needs to be at least 0.03% to be detected in a sample of 10,000 sequences with a chance of 95%. This corresponds to more than a million cells per gram of soil. However, many soil bacteria will have a lower abundance, at least temporarily or locally. A second important question is that of how well the dataset from pyrosequencing can represent the community structure. The simulation results in Table 2.1.3 show that there will be a significant error in predicting the relative abundance of rare species. A relative species abundance of less than 0.01% will be predicted with an error of more than 100%. Obviously, it is still very difficult to reliably predict the community structure also with respect to minorities.

This technique was also applied to study the soils samples in both experiments. Previously discussed advantages and limitations for this technique were nicely confirmed.

1.3.4 Methods to study the function of soil microbial community

1.3.4.1 Targeted detection of functional genes (chapter 2)

PCR-based detection of functional genes in soil has widely been used in several studies. The presence of different functional genes in soil bacteria has usually been studied by PCR amplification from total DNA and subsequent cloning and sequencing or high throughput sequencing. If sufficient amounts of PCR products can be obtained, the analysis of PCR amplicons by cloning and sequencing or by molecular fingerprints can provide information on the diversity of respective genes, e.g., *pmoA*, [43], bacterial or archaeal *amoA*[44,45], *ndo* genes[12,13,16,18,46], or chitinase genes [47].

If the gene targeted is present above a critical detection threshold then its presence can be quantified by real-time PCR [48,49,50]. In particular, the development of real-time PCR assays enabled studies relating gene abundance to various different biotic or abiotic factors. Again the detection limit is an important issue. Only if gene copies occur with more than 10^5-10^6 copies per gram of soil, amplicons visible in ethidium bromide-stained agarose gels are obtained. The detection limit is comparable to real-time PCR. Although SybrGreen detection is widely employed to detect functional genes by quantitative PCR, the use of TaqMan probes might increase specificity and sensitivity of the system. In many studies the copy number determined is related to gram of soil or even µl of DNA. However, it is strongly recommended that the copy number determined is related to 16S rRNA gene copy numbers to overcome biases caused by uncertainties of the total soil DNA composition. Although comparable amounts of DNA are obtained per gram of soil, the number of 16S rRNA gene copy numbers might vary considerably [51]. If the gene targeted is less abundant in the soil population then it is recommendable to increase the sensitivity of their detection by Southern blot hybridization or by nested or semi-nested PCR approaches. Even if amplicons are not visible on the agarose gel they can be detected by means of hybridization with digoxigenin or ^{32}P labeled probes. Although most of the data obtained are semi-quantitative the sensitivity of gene detection is often increased by two orders of magnitude. In addition, by choosing the stringency of the hybridization conditions sequence specificity of the signal can be confirmed. The presence of different functional genes in soil bacteria

has been studied by PCR amplification from total DNA and subsequent analysis by Southern blot analysis, e.g., *merA*, *trfA [52]*, *nahA*, *nahH, phnAC [12,15,18]*. The spectrum of primers is also a very important issue as the diversity of many genes is still unknown and many novel sequences could be frequently found from environmental samples.

1.3.4.2 Non-targeted detection of genes

1.3.4.2.1 Functional microarray (chapter 2)

Functional genes have been detected in soil DNA by means of so-called functional arrays. However, the performance of microarray hybridization from complex environments such as soil has to be carefully evaluated and obviously a number of technical challenges need to be overcome before this technique can be fully exploited. A microarray with 100 functional genes was used by Wu *et al*. [54] to systematically study the specificity, sensitivity and quantification of microarray hybridization with DNA from complex environmental samples. When environmental DNA is used without prior PCR amplification this seems to be the most difficult challenge. The level of detection is 1,000 to 10,000 fold lower than with PCR amplification [54]. In order to achieve a comprehensive investigation of functional genes, so-called GeoChip microarrays containing over 24,000 probes covering more than 10,000 genes distributed among more than 150 functional groups involved in nitrogen, carbon, sulphur and phosphorus cycling were recently employed for soil studies [55,56,57]. The direct hybridizations were, however, not very sensitive, and to increase sensitivity of the approach, a pre-amplification step (e.g. using rolling circle amplification) is now included. Despite numerous challenges that need to be solved, the GeoChip is a powerful tool that can be used to link functional microbial dynamics to particular ecosystem processes. Probe development, hybridization quality and data evaluation are crucial steps for an appropriate use of DNA microarrays to study the soil microbiota. The GeoChip was applied to study N- and C-cycle genes in DNA extracted from soil taken from different sites in the Antarctic [58]. The N- and C-cycle genes detected differed significantly across sampling locations and vegetation types. However, DNA microarrays cannot generate information on new sequence types and thus only the breadth of known functions can be assessed [58]. A recent study of Suenaga *et al.* gives an idea of the severity of this limitation [59]. They created a

metagenomic library with DNA extracted from activated sludge and sequenced those clones with extradiol dioxygenase activity. By sequence analysis, they found that most of these clones (36 of 38) contained complete aromatic degradation pathways that shared low similarity to those found in known cultured bacteria metabolizing aromatic compounds. Seeing that soil bacterial communities can be extremely diverse, it will be a challenge to include all potentially important genetic variants of a particular soil function on a microarray. Furthermore, the abundance of most bacteria may be below the detection limit of microarrays, some of them having the potential to take over important functions when environmental conditions are advantageous for them. To increase the sensitivity, community DNA was frequently subjected to whole genome amplification before microarray analysis. During a stay in the laboratory of J. Zhou, University of Oklahoma we also applied the GeoChip 2 to study the changes of functional genes after phenanthrene spiking (see chapter 4).

1.3.4.2.2 High throughput sequencing of functional genes

High throughput sequencing technologies such as Roche 454, Illumina and ABI SOLiD, not only improve our understanding of the structure of microbial communities but also extend the insight into their function. By high throughput sequencing of cDNA library, detailed information on transcription of complex microbial community could also be acquired. Practically, the majority of the RNA pool is rRNA and tRNA, while the mRNA is only a small fraction [60]. Therefore, to study the transcriptome, enrichment of mRNA is frequently needed. Because of the poly-A tail, mRNA of eukaryotes could be easily enriched by hybridization with poly-T region [61]. Even though poly-A is lacking, so far there are still some methods to specifically enrich bacterial mRNA [62,63]. Frequently, direct sequencing of the cDNA pool could also have an advantage as both the microbial community structure and functions could be studied [64]. Using high throughput sequencing of cDNA could supply information of potential active genes or bacteria. But due to the short half-time and instability of RNA, these methods might be unsuitable to study long-term effects on microbial communities.

1.3.4.2.3 Metagenomic libraries (chapter 2)

Another non-targeted approach is cloning large fragments of environmental DNA into fosmid or BAC vectors and subsequent screening of the metagenomic libraries.

Metagenomics offers access to functional genes in the soil microbiota and thus to novel bioactive products. This approach has been successfully applied to recover DNA coding for so far unknown enzymes or antibiotics directly from soil DNA [65,66]. However, despite the vast potential of this approach, several methodological challenges remain to be solved. To obtain sequence information of particular uncultured taxa which are assumed to be abundant in the soil sample, in a first step specific primers are used to identify the clones carrying the respective genes, either based on 16S rRNA or even on functional genes. Using this approach, genomic fragments of uncultured *Acidobacteria* or *Archaea* could be obtained and subjected to sequence-based gene detection [67,68,69,70]. Sequencing of a large fosmid library from a grassland soil revealed that a fragment with a rRNA gene, which showed the highest similarity to group 1.1b of the crenarchaeota, carried genes encoding copper nitrite reductase and two subunits of ammonia monooxygenase or particulate methane monooxygenase [68]. The discovery of ammonia oxidizing *Archaea* showed the potential of soil metagenomic studies to elucidate the role of uncultured organisms in, e.g., soil nutrient cycles [71]. Considering that the majority of soil bacteria occurs at relatively low abundance but might have an important ecological role, the development of approaches such as targeted access to low-abundance bacteria are of considerable interest. Recently Hjort *et al.* (2010) analyzed chitinase genes in DNA directly extracted from a phytopathogen-suppressive soil, in a metagenomic library constructed from microbial cells extracted from soil, and in genomic DNA from bacterial isolates with antifungal and chitinase activities [72]. Although T-RFLP of chitinase genes revealed differences in amplified chitinase genes from the metagenomic library and the directly extracted DNA, approximately 40% of the identified chitinase T-RFs were found in both sources. All of the chitinase T-RFs from the isolates were matched to T-RFs in the directly extracted DNA and in the metagenomic library. The authors demonstrated an impressive agreement between three very different screening techniques all of which pointed towards specific *Streptomyces* species that could play a role in suppression of fungi by chitinase production in soil. However, some clusters of chitinase genes were represented depending on the approach used. Therefore, Hjort *et al.* (2010) concluded that a combination of molecular approaches increases the information obtained and the reliability of the data.

1.3.4.2.4 Cultivation-independent detection of genes carried by mobile genetic elements (chapter 2)

The importance of horizontal gene exchange for short-term bacterial adaptability and for successful colonization of new ecological niches has only recently been appreciated [73]. Nowadays mobile genetic elements (MGE) are recognized as an important and essential component that promotes bacterial diversity. The PCR-based detection of MGE-specific sequences in community DNA was first used by Götz *et al.* [74]. Primers targeting replicon-specific sequences were designed on the basis of sequenced broad host range plasmids or integrons. In combination with Southern blot hybridization a specific and sensitive monitoring of large numbers of environmental samples became possible [51,52,75]. While the targeted detection of MGE by endpoint or real-time PCR requires sequence information, capturing MGE directly from soil microbial pellets and subsequent sequencing allows a non-targeted discovery and detection of genes. Direct capturing of MGE by means of so-called exogenous plasmid isolation techniques [76,77] that has been widely used to capture MGE conferring selectable traits such as mercury or antibiotic resistance have been acquired from a wide range of soil or sediments in gram-negative recipients functioning as a genetic sink [51,52,75]. Capturing of degradative genes resident on MGE has been demonstrated as well [78]. Increased transfer frequencies have often been observed when the soil environmental sample was previously exposed to pollutants. The PCR-based detection of MGE has been used for monitoring the abundance of MGE in soils [51,74,75] and “hot spots” with high abundance of MGE could be identified. In our study, a PCR targeting IncP 9 plasmids which harbor gene clusters for PAH mineralization was developed and applied (chapter 5). Due to the absence of responses of IncP 9 plasmids carrying bacterial populations to phenanthrene spiking, efforts to exogenously isolate IncP-9 plasmids were not made.

1.4 The interactions between soil microbes and organic pollutants

The fate of persistent organic pollutants (POPs) in soil is of great interest due to their potential for bioaccumulation, persistence, transport and toxicity [79,80,81]. Several factors including the soil type (mineral and organic matter, microbial communities present) and the physicochemical properties (molecular structure, hydrophobicity, water solubility) control the fate and behavior of POPs in the soil system. The putative fate and behavior of POPs such as phenanthrene in soil are

biodegradation, bioaccumulation, sequestration, leaching and volatilization [82]. When POPs enter the porous soil system they first contaminate macropores followed by drainage of successively smaller pores [83]. POPs can reach also extremely small pores (< 1 µm) where the pollutants appear to be inaccessible to biodegradation because microbes will be largely absent in very small pores. Mass transfer of POPs is required before they can be taken up and degraded [84]. Furthermore, it was shown that the biodegradative potential varies between different soils and the soil organic matter has an important influence on the fate of POPs [85,86].

1.4.1 Primers specifically detecting key PAH degrading genes (chapter2)

The potential for microbial degradation of polycyclic aromatic hydrocarbons (PAH) in the environment is of major interest due to the bioaccumulation and toxicity of these compounds. Key degrading genes were used to evaluate the PAH degrading capacity of soil microbes. How the diversity and abundance are influenced by different soil types or their history and type of pollution is not yet well explored, because the existing primer systems have severe limitations. In Table 2.1.4, published primers that detect PAH ring hydroxylating dioxygenase genes are summarized. Most of the primers target only a rather narrow range of sequences, e.g., *nahAc* or *phnAc* type sequences, or only genes from gram-negative bacteria. One primer system could detect dioxygenase genes of both Gram-positive and Gram-negative bacteria, but the specificity is broader than PAH dioxygenases, and the short amplified fragment does not resolve much of the diversity [53]. To overcome these limitations, recently accumulating sequence data need to be used to develop better primer systems to determine the diversity and abundance of PAH ring hydroxylating dioxygenase genes of soil microbial communities. A novel PCR targeting at *PAH-RHDα* genes was developed and applied for testing the hypothesis of soil type dependent responses of bacteria carrying *PAH-RH*Dα genes to phenanthrene spiking (see chapter 3).

1.4.2 Important taxonomic groups responding to PAH contaminations

A wide range of taxonomic groups of bacteria are known to be capable to utilize different PAHs as sole carbon source. Their ecological fitness was intensively studied by molecular fingerprinting. Frequently, the aspects of interests are shifts in bacterial communities after PAH perturbations, the survival of bio-degraders in soil or other complex systems and the microbial community structure in contaminated soils. By PCR-DGGE, Gomes et al. (2005) found that *Pseudomonas putida* KT2442 could propagate in naphthalene-polluted soils. *Burkholderia* sp. RP007 and *Nocardia vinaceae* like bacteria were found enriched only in naphthalene-polluted soils without inoculation of *P. putida* KT2442 [15]. Using nested PCR-DGGE approach, Leys et al. (2005) found that in contaminated soils mycobacterial communities cannot only be distinguished with known PAH-degrading *Mycobacteria* species such as *M. frederiksbergense* and *M. austroafricanum*, but also with *M. tusciae* which was not reported to be involved in PAH mineralization [87]. Clone library was also frequently applied together to study the microbial community structure. Bordenave et al. (2007) studied the succession of the bacterial community in heavy oil contaminated soils by T-RFLP and clone library studies of 16S rRNA gene. First, *Pseudomonas* and *Acinetobacter* responded and became dominant and then were replaced by "*Bacillus*" and "*Alphaproteobacteria*" [88].

1.4.3 Applications of functional gene arrays and metagenomic library

Unlike the PCR-clone library or PCR-fingerprinting for one kind of genes, functional gene arrays could be used to study the comprehensive functions which related with thousands of genes. A global map of the functional composition of complex microbial communities could theoretically be acquired. Leigh et al. (2007) studied the ^{13}C-DNA pool from microcosm with ^{13}C-labeled biphenyl as contaminants by GeoChip [89]. Except for bacterial aromatic hydrocarbon degradation genes, plant aromatic catabolic pathway genes in ß-ketoadipate pathway were also detected. It is also interesting to notice that eight out of 27 aromatic hydrocarbon degradation genes were detected from total community DNA. This might be an indication that not only the presence but also the relative abundance of certain genes decide whether it could be detected. Another successful application of GeoChip is the study performed

by Wang et al. [90], in which the diversity of functional genes of deep sea hydrothermal vents was studied. Their results indicate highly functional diversity of hydrothermal microbial community, which responds rapidly to temperature and chemical gradients. In both studies, the data were explained as present and absent but not the signal intensity of each gene. This might be due to the unevenness of RCA amplification. Unlike other methods, metagenomic library provide a chance not only to check single genes but also the pathway. Suenaga et al. [91] cloned the community DNA extracted from activate sludge into fosmids. These fosmids with extradiol dioxygenase activity were subjected to shotgun sequencing. It was found that not only many genes shared low similarity with the corresponding well known ones but also the majority contigs encoding partial pathways for PAHs mineralization. This finding suggested that the degradation of aromatic hydrocarbon could happen through integration of fragmental pathways [91]. Due to that relative low numbers of fosmid clones could be analyzed, only a small fraction of bacteria in soil could be studied. In addition, the bias caused by the host strains is another limitation.

1.4.4 Mobile genetic elements (MGE)

Plasmids are important MGE which contribute to the adaption of *Bacteria* to changing environments. For example, plasmids belonging to the IncP-9 group are known to frequently carry the genetic pathway for PAH mineralization [31]. Three out of four completely sequenced plasmids NAH7, pDTG1 and pNAH20 bear upper operon for naphthalene degradation which encodes enzymes catalyzing the transformation of naphthalene to salicylate under aerobic condition. By conjugative transfer, this group of plasmids can shuttle among a wide spectrum of *Pseudomonas* species. Their geographical distribution is also extremely wide, ranging from Asia, Europe to North America, and is even found in oil-contaminated Antarctic soil [92]. And also their fitted niches are very diverse from coal-tar-contaminated soil, oil-contaminated soil, to manure exposure to antibiotics. It was assumed that IncP-9 plasmids contribute globally to the adaption of *Pseudomonas* species to polluted environments by aromatic compounds, heavy metal, antibiotics and so on. Ma *et al.* isolated 22 PAH degrading *Pseudomonas* from contaminated Antarctic soils [93]. Further studies with Southern blot found that corresponding *ndo* genes were located on plasmids. PAH-degrading plasmids were also found in *Sphingomonas*. Cho and

Kim found that *Sphingomonas* sp. Strain KS14 harbors a mega plasmid which could be responsible for the mineralization of phenanthrene and naphthalene [94].

1.5 Outlook: Biogeochemical interfaces in soils

So far, most microbial ecological studies mainly focus on two questions: the first one is being “who is there?” the second “what are they doing?” Studies were frequently performed at relative macro-scale (~cm, gram). The structure of soil bacterial community was found influenced by several factors, such as soil type, plant covers, animal activities, contamination type and levels. In most studies, the community structure within the same treatments was found to be highly similar. The development of a soil microbial community seems not to be chaotic even though the initial community structure is one of essential determinants. Modelling the process and microbial community structure still seems possible. But the work is extremely difficult because on the one side huge amounts of bacteria with an extreme diversity are present in each soil; on the other side the links between microbes and microbe driven biogeochemical processes are still largely unknown. Micro-scale studies are believed to be one possible key to unlock these hidden secrets and eventually provide essential data for the modelling.

Soils, no matter of which texture, are a very complex heterogeneous matrix in which gradients of chemicals, air conditions and microbes form “hot spots”. Several materials such as organic matter, minerals, and iron oxides are believed to be major components in controlling the formation of interfaces in soil. Therefore the mineral composition could be key forces which structure the microbial community. To test this hypothesis, artificial soils of different mineral composition were inoculated with the same microbial inoculants. And the microbial communities were monitored at different sampling time. Details were described in chapter 7.

1.5.1 Absorption/desorption of charcoal, clay minerals and iron oxides to pollutants.

Charcoal is of microporous structure and thus has a very large surface area, which enables it to adsorb chemicals and form unique environmental niches e.g. forest fire sites. The influence of charcoal in absorbing polycyclic aromatic hydrocarbons was explored in several studies. In the studies of Ahn *et al.*, charcoal was found with selective adsorption ability towards phenanthrene to surfactant TX100 [95]. This kind

of selectivity is influenced by the size of particles, the bigger the particles, the stronger the selectivity [95]. But in another study by Lebo *et al.*, fine charcoal was found to be more efficient in blocking phenanthrene than coarse charcoal [96]. Despite the size of the particles, the aromaticity of charcoal was also found to positively correlate with its ability in absorption of organic pollutants. Sun *et al.* suggested that the presence pi–pi electron donor-acceptor interaction might be the molecular mechanism [97]. The absorption of phenanthrene by charcoal is reversible reaction. The rate of phenanthrene desorption from charcoal showed no correlation with particle size. All three tested charcoals with varied particle size are of comparable desorption rate for phenanthrene [98]. The absorption and desorption of phenanthrene by charcoal could also influence its bioavailability, mineralization, transporting and so on. Reduced mineralization rate and extent for phenanthrene were observed. The higher the concentration of charcoal in soil, the stronger the reduction of mibneralization rates [99]. Clay minerals and iron oxides such as smectites, illite, montmorillonite, and kaolinite also play a role in transporting and bioavailability of aromatic compounds in soil [100]. Their sorption ability is influenced by many factors such as pH, iron strength and temperature [101]. Exchangeable cations were also found to relate with absorption ability of clays. In the study of Zhang *et al.* clays have a significantly elevated sorption ability to 2, 4-dinitrotoluene and nitrobenzene after saturating with K^+ and Cs^+ [102].

1.5.2 Influence of charcoal, clay minerals and iron oxides on the function and structure of soil microbial community.

The influence of physicochemical characters on the function and structure of microbial communities was explored in many studies. Several parameters such as pH, salinity, temperature, pore size were reported as the driving force of soil bacterial diversity [20,21,23,103]. Recently, the influences of charcoal on functions of soil microbes were explored. Increased abundance of nitrification and ammonium-oxidizing bacteria were observed in charcoal rich soils by Ball et al. [104]. Leglize et al. showed that three biofilm-forming bacteria showed improved phenanthrene degrading ability when charcoal was added to the system [105]. But the influences of charcoal, clay minerals and iron oxides on the soil microbial communities were rarely studied.

1. Standing D, Baggs EM, Wattenbach M, Smith P, Killham K (2007) Meeting the challenge of scaling up processes in the plant-soil-microbe system. Biology and Fertility of Soils 44: 245-257.

2. Smalla K, Wieland G, Buchner A, Zock A, Parzy J, et al. (2001) Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: Plant-dependent enrichment and seasonal shifts revealed. Applied and Environmental Microbiology 67: 4742-4751.
3. Nihorimbere V, Fickers P, Thonart P, Ongena M (2009) Ecological fitness of Bacillus subtilis BGS3 regarding production of the surfactin lipopeptide in the rhizosphere. Environmental Microbiology Reports 1: 124-130.
4. Coelho MRR, Marriel IE, Jenkins SN, Lanyon CV, Seldin L, et al. (2009) Molecular detection and quantification of nifH gene sequences in the rhizosphere of sorghum (Sorghum bicolor) sown with two levels of nitrogen fertilizer. Applied Soil Ecology 42: 48-53.
5. Child R, Miller CD, Liang Y, Sims RC, Anderson AJ (2007) Pyrene mineralization by Mycobacterium sp strain KMS in a Barley rhizosphere. Journal of Environmental Quality 36: 1260-1265.
6. Piutti S, Hallet S, Rousseaux S, Philippot L, Soulas G, et al. (2002) Accelerated mineralisation of atrazine in maize rhizosphere soil. Biology and Fertility of Soils 36: 434-441.
7. Nechitaylo TY, Yakimov MM, Godinho M, Timmis KN, Belogolova E, et al. (2010) Effect of the earthworms Lumbricus terrestris and Aporrectodea caliginosa on bacterial diversity in Soil. Microbial Ecology 59: 574-587.
8. Gomez-Brandon M, Lazcano C, Lores M, Dominguez J (2010) Detritivorous earthworms modify microbial community structure and accelerate plant residue decomposition. Applied Soil Ecology 44: 237-244.
9. Hery M, Singer AC, Kumaresan D, Bodrossy L, Stralis-Pavese N, et al. (2008) Effect of earthworms on the community structure of active methanotrophic bacteria in a landfill cover soil. ISME Journal 2: 92-104.
10. Marschner P, Crowley D, Yang CH (2004) Development of specific rhizosphere bacterial communities in relation to plant species, nutrition and soil type. Plant and Soil 261: 199-208.
11. Marschner P, Yang CH, Lieberei R, Crowley DE (2001) Soil and plant specific effects on bacterial community composition in the rhizosphere. Soil Biology & Biochemistry 33: 1437-1445.
12. Ding GC, Heuer H, Zuhlke S, Spiteller M, Pronk GJ, et al. (2010) Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes by using a novel PCR detection system. Appl Environ Microbiol 76: 4765-4771.
13. Flocco CG, Gomes NC, Mac CW, Smalla K (2009) Occurrence and diversity of naphthalene dioxygenase genes in soil microbial communities from the Maritime Antarctic. EnvironMicrobiol 11: 700-714.
14. Kim YH, Engesser KH, Cerniglia CE (2005) Numerical and genetic analysis of polycyclic aromatic hydrocarbon-degrading mycobacteria. MicrobEcol 50: 110-119.
15. Gomes NC, Kosheleva IA, Abraham WR, Smalla K (2005) Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. FEMS MicrobiolEcol 54: 21-33.

16. Bordenave S, Goni-Urriza M, Vilette C, Blanchard S, Caumette P, et al. (2008) Diversity of ring-hydroxylating dioxygenases in pristine and oil contaminated microbial mats at genomic and transcriptomic levels. EnvironMicrobiol 10: 3201-3211.
17. Schafer AN, Snape I, Siciliano SD (2009) Influence of liquid water and soil temperature on petroleum hydrocarbon toxicity in Antarctic soil. Environ Toxicol Chem 28: 1409-1415.
18. Gomes NC, Borges LR, Paranhos R, Pinto FN, Krögerrecklenfort E, et al. (2007) Diversity of *ndo* genes in mangrove sediments exposed to different sources of polycyclic aromatic hydrocarbon pollution. ApplEnvironMicrobiol 73: 7392-7399.
19. Johnsen AR, Karlson U (2005) PAH degradation capacity of soil microbial communities—does it depend on PAH exposure? Microb Ecol 50: 488-495.
20. Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. Proc Natl Acad Sci U S A 104: 11436-11440.
21. Fierer N, Jackson RB (2006) The diversity and biogeography of soil bacterial communities. Proc Natl Acad Sci U S A 103: 626-631.
22. Hollister EB, Engledow AS, Hammett AJ, Provin TL, Wilkinson HH, et al. (2010) Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments. ISME J 4: 829-838.
23. Carson JK, Gonzalez-Quinones V, Murphy DV, Hinz C, Shaw JA, et al. (2010) Low pore connectivity increases bacterial diversity in soil. Appl Environ Microbiol 76: 3936-3942.
24. Rentz JA, Alvarez PJ, Schnoor JL (2008) Benzo[a]pyrene degradation by Sphingomonas yanoikuyae JAR02. Environ Pollut 151: 669-677.
25. Kim SJ, Kweon O, Jones RC, Freeman JP, Edmondson RD, et al. (2007) Complete and integrated pyrene degradation pathway in Mycobacterium vanbaalenii PYR-1 based on systems biology. J Bacteriol 189: 464-472.
26. Yagi JM, Sims D, Brettin T, Bruce D, Madsen EL (2009) The genome of *Polaromonas naphthalenivorans* strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. EnvironMicrobiol 11: 2253-2270.
27. Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. MicrobiolRev 59: 143-169.
28. Hutchison CA, 3rd, Venter JC (2006) Single-cell genomics. Nat Biotechnol 24: 657-658.
29. Daniel R (2005) The metagenomics of soil. Nat Rev Microbiol 3: 470-478.
30. Hultman J, Auvinen P (2010) [Metagenomics opens up new frontiers in microbiology]. Duodecim 126: 1278-1285.
31. Sota M, Yano H, Ono A, Miyazaki R, Ishii H, et al. (2006) Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase. J Bacteriol 188: 4057-4067.
32. Stuart-Keil KG, Hohnstock AM, Drees KP, Herrick JB, Madsen EL (1998) Plasmids responsible for horizontal transfer of naphthalene catabolism genes between bacteria at a coal tar contaminated site are homologous to pDTG1 from Pseudomonas putida NCIB 9816-4. Appl. Environ. Microbiol. 64: 3633-3640.
33. Tebbe CC, Wenderoth DF, Vahjen W, Lubke K, Munch JC (1995) Direct detection of recombinant gene expression by two genetically engineered yeasts in soil on the transcriptional and translational levels. Appl Environ Microbiol 61: 4296-4303.

34. Wang Y, Qian PY (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PloS One 4: e7401.
35. Brodie EL, Desantis TZ, Joyner DC, Baek SM, Larsen JT, et al. (2006) Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Appl Environ Microbiol 72: 6288-6298.
36. DeSantis TZ, Brodie EL, Moberg JP, Zubieta IX, Piceno YM, et al. (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microb Ecol 53: 371-383.
37. DeAngelis KM, Brodie EL, DeSantis TZ, Andersen GL, Lindow SE, et al. (2009) Selective progressive response of soil microbial community to wild oat roots. ISME J 3: 168-178.
38. Campbell BJ, Polson SW, Hanson TE, Mack MC, Schuur EA (2010) The effect of nutrient deposition on bacterial communities in Arctic tundra soil. Environ. Microbiol.
39. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537-7541.
40. Schloss PD (2010) The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. PloS Comput Biol 6: e1000844.
41. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73: 5261-5267.
42. Schloss PD, Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. ApplEnvironMicrobiol 71: 1501-1506.
43. Horz HP, Yimga MT, Liesack W (2001) Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of pmoA, mmoX, mxaF, and 16S rRNA and ribosomal DNA, including pmoA-based terminal restriction fragment length polymorphism profiling. Appl Environ Microbiol 67: 4177-4185.
44. Jia Z, Conrad R (2009) Bacteria rather than Archaea dominate microbial ammonia oxidation in an agricultural soil. Environ Microbiol 11: 1658-1671.
45. Nicol GW, Leininger S, Schleper C, Prosser JI (2008) The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. Environ Microbiol 10: 2966-2978.
46. Yagi JM, Madsen EL (2009) Diversity, abundance, and consistency of microbial oxygenase expression and biodegradation in a shallow contaminated aquifer. ApplEnvironMicrobiol 75: 6478-6487.
47. Enkerli J, Ghormade V, Oulevey C, Widmer F (2009) PCR-RFLP analysis of chitinase genes enables efficient genotyping of Metarhizium anisopliae var. anisopliae. J Invertebr Pathol 102: 185-188.
48. Schauss K, Focks A, Leininger S, Kotzerke A, Heuer H, et al. (2009) Dynamics and functional relevance of ammonia-oxidizing archaea in two agricultural soils. Environ Microbiol 11: 446-456.
49. Hai B, Diallo NH, Sall S, Haesler F, Schauss K, et al. (2009) Quantification of key genes steering the microbial nitrogen cycle in the rhizosphere of sorghum cultivars in tropical agroecosystems. Appl Environ Microbiol 75: 4993-5000.

50. Heuer H, Smalla K (2007) Manure and sulfadiazine synergistically increased bacterial antibiotic resistance in soil over at least two months. Environ Microbiol 9: 657-666.
51. Heuer H, Kopmann C, Binh CT, Top EM, Smalla K (2009) Spreading antibiotic resistance through spread manure: characteristics of a novel plasmid type with low %G+C content. Environ Microbiol 11: 937-949.
52. Smalla K, Haines AS, Jones K, Krögerrecklenfort E, Heuer H, et al. (2006) Increased abundance of IncP-1beta plasmids and mercury resistance genes in mercury-polluted river sediments: first discovery of IncP-1beta plasmids with a complex mer transposon as the sole accessory element. Appl Environ Microbiol 72: 7253-7259.
53. Ni Chadhain SM, Norman RS, Pesce KV, Kukor JJ, Zylstra GJ (2006) Microbial dioxygenase gene population shifts during polycyclic aromatic hydrocarbon biodegradation. ApplEnvironMicrobiol 72: 4078-4087.
54. Wu L, Thompson DK, Li G, Hurt RA, Tiedje JM, et al. (2001) Development and evaluation of functional gene arrays for detection of selected genes in the environment. Appl Environ Microbiol 67: 5780-5790.
55. Wu L, Liu X, Schadt CW, Zhou J (2006) Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl Environ Microbiol 72: 4931-4941.
56. He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, et al. (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J 1: 67-77.
57. Gentry TJ, Wickham GS, Schadt CW, He Z, Zhou J (2006) Microarray applications in microbial ecology research. Microb Ecol 52: 159-175.
58. Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA (2007) Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. ISME J 1: 163-179.
59. Suenaga H, Koyama Y, Miyakoshi M, Miyazaki R, Yano H, et al. (2009) Novel organization of aromatic degradation pathway genes in a microbial community as revealed by metagenomic analysis. ISME J 3: 1335-1348.
60. Stewart FJ, Ottesen EA, DeLong EF (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. ISME J 4: 896-907.
61. Kuribayashi-Ohta K, Tamatsukuri S, Hikata M, Miyamoto C, Furuichi Y (1993) Application of oligo(dT)30-latex for rapid purification of poly(A)+ mRNA and for hybrid subtraction with the in situ reverse transcribed cDNA. Biochim Biophys Acta 1156: 204-212.
62. Mettel C, Kim Y, Shrestha PM, Liesack W (2010) Extraction of mRNA from soil. Appl Environ Microbiol 76: 5995-6000.
63. Gilbert JA, Field D, Huang Y, Edwards R, Li W, et al. (2008) Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. PloS One 3: e3042.
64. Urich T, Lanzen A, Qi J, Huson DH, Schleper C, et al. (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. PloS One 3: e2527.
65. Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, et al. (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl Environ Microbiol 66: 2541-2547.

66. Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, et al. (2002) Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. Appl Environ Microbiol 68: 4301-4306.
67. Ochsenreiter T, Selezi D, Quaiser A, Bonch-Osmolovskaya L, Schleper C (2003) Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. Environ Microbiol 5: 787-797.
68. Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, et al. (2005) Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. Environ Microbiol 7: 1985-1995.
69. van Elsas JD, Costa R, Jansson J, Sjoling S, Bailey M, et al. (2008) The metagenomics of disease-suppressive soils — experiences from the METACONTROL project. Trends Biotechnol 26: 591-601.
70. van Elsas JD, Speksnijder AJ, van Overbeek LS (2008) A procedure for the metagenomics exploration of disease-suppressive soils. J Microbiol Methods 75: 515-522.
71. Leininger S, Urich T, Schloter M, Schwark L, Qi J, et al. (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442: 806-809.
72. Hjort K, Bergstrom M, Adesina MF, Jansson JK, Smalla K, et al. (2010) Chitinase genes revealed and compared in bacterial isolates, DNA extracts and a metagenomic library from a phytopathogen-suppressive soil. FEMS Microbiol Ecol 71: 197-207.
73. Heuer H, Abdo Z, Smalla K (2008) Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. FEMS Microbiol Ecol 65: 361-371.
74. Götz A, Pukall R, Smit E, Tietze E, Prager R, et al. (1996) Detection and characterization of broad-host-range plasmids in environmental bacteria by PCR. Appl Environ Microbiol 62: 2621-2628.
75. Heuer H, Krögerrecklenfort E, Wellington EM, Egan S, Elsas JD, et al. (2002) Gentamicin resistance genes in environmental bacteria: prevalence and transfer. FEMS Microbiol Ecol 42: 289-302.
76. Bale MJ, Day MJ, Fry JC (1988) Novel method for studying plasmid transfer in undisturbed river epilithon. Appl Environ Microbiol 54: 2756-2758.
77. Hill KE, Weightman AJ, Fry JC (1992) Isolation and screening of plasmids from the epilithon which mobilize recombinant plasmid pD10. Appl Environ Microbiol 58: 1292-1300.
78. Top EM, Holben WE, Forney LJ (1995) Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl Environ Microbiol 61: 1691-1698.
79. Habe H, Omori T (2003) Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. Bioscience Biotechnology and Biochemistry 67: 225-243.
80. Johnsen AR, Wick LY, Harms H (2005) Principles of microbial PAH-degradation in soil. EnvironPollut 133: 71-84.
81. Kanaly RA, Harayama S (2000) Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. J. Bacteriol. 182: 2059-2067.
82. Semple KT, Morriss AWJ, Paton GI (2003) Bioavailability of hydrophobic organic contaminants in soils: fundamental concepts and techniques for analysis. European Journal of Soil Science 54: 809-818.

83. Nam K, Alexander M (1998) Role of nanoporosity and hydrophobicity in sequestration and bioavailability: Tests with model solids. Environmental Science & Technology 32: 71-74.
84. Bosma TNP, Middeldorp PJM, Schraa G, Zehnder AJB (1997) Mass transfer limitation of biotransformation: Quantifying bioavailability. Environmental Science & Technology 31: 248-252.
85. Macleod CJ, Semple KT (2002) The adaptation of two similar soils to pyrene catabolism. Environ Pollut 119: 357-364.
86. Denef K, Six J (2005) Clay mineralogy determines the importance of biological versus abiotic processes for macroaggregate formation and stabilization. European Journal of Soil Science 56: 469-479.
87. Leys NM, Ryngaert A, Bastiaens L, Wattiau P, Top EM, et al. (2005) Occurrence and community composition of fast-growing Mycobacterium in soils contaminated with polycyclic aromatic hydrocarbons. FEMS Microbiol Ecol 51: 375-388.
88. Bordenave S, Goni-Urriza MS, Caumette P, Duran R (2007) Effects of heavy fuel oil on the bacterial community structure of a pristine microbial mat. Appl Environ Microbiol 73: 6089-6097.
89. Leigh MB, Pellizari VH, Uhlik O, Sutka R, Rodrigues J, et al. (2007) Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). ISME J 1: 134-148.
90. Wang F, Zhou H, Meng J, Peng X, Jiang L, et al. (2009) GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de Fuca Ridge hydrothermal vent. Proc Natl Acad Sci U S A 106: 4840-4845.
91. Suenaga H, Koyama Y, Miyakoshi M, Miyazaki R, Yano H, et al. (2009) Novel organization of aromatic degradation pathway genes in a microbial community as revealed by metagenomic analysis. ISME Journal 3: 1335-1348.
92. Sevastsyanovich YR, Krasowiak R, Bingle LE, Haines AS, Sokolov SL, et al. (2008) Diversity of IncP-9 plasmids of Pseudomonas. Microbiology 154: 2929-2941.
93. Ma Y, Wang L, Shao Z (2006) Pseudomonas, the dominant polycyclic aromatic hydrocarbon-degrading bacteria isolated from Antarctic soils and the role of large plasmids in horizontal gene transfer. Environ Microbiol 8: 455-465.
94. Cho JC, Kim SJ (2001) Detection of mega plasmid from polycyclic aromatic hydrocarbon-degrading Sphingomonas sp. Strain KS14. J Mol Microbiol Biotechnol 3: 503-506.
95. Ahn CK, Kim YM, Woo SH, Park JM (2007) Selective adsorption of phenanthrene dissolved in surfactant solution using activated carbon. Chemosphere 69: 1681-1688.
96. Lebo JA, Huckins JN, Petty JD, Cranor WL, Ho KT (2003) Comparisons of coarse and fine versions of two carbons for reducing the bioavailabilities of sediment-bound hydrophobic organic contaminants. Chemosphere 50: 1309-1317.
97. Sun H, Zhou Z (2008) Impacts of charcoal characteristics on sorption of polycyclic aromatic hydrocarbons. Chemosphere 71: 2113-2120.
98. Morelis S, van Noort PC (2008) Kinetics of phenanthrene desorption from activated carbons to water. Chemosphere 71: 2044-2049.
99. Rhodes AH, McAllister LE, Chen R, Semple KT (2010) Impact of activated charcoal on the mineralisation of 14C-phenanthrene in soils. Chemosphere 79: 463-469.

100. Liu C, Li H, Teppen BJ, Johnston CT, Boyd SA (2009) Mechanisms associated with the high adsorption of dibenzo-p-dioxin from water by smectite clays. Environ Sci Technol 43: 2777-2783.
101. Peng X, Wang J, Fan B, Luan Z (2009) Sorption of endrin to montmorillonite and kaolinite clays. J Hazard Mater 168: 210-214.
102. Zhang D, Zhu D, Chen W (2009) Sorption of nitroaromatics to soils: comparison of the importance of soil organic matter versus clay. Environ Toxicol Chem 28: 1447-1454.
103. Miller SR, Strong AL, Jones KL, Ungerer MC (2009) Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. Appl Environ Microbiol 75: 4565-4572.
104. Ball PN, MacKenzie MD, DeLuca TH, Holben WE (2010) Wildfire and charcoal enhance nitrification and ammonium-oxidizing bacterial abundance in dry Montane forest soils. J Environ Qual 39: 1243-1253.
105. Leglize P, Alain S, Jacques B, Corinne L (2008) Adsorption of phenanthrene on activated carbon increases mineralization rate by specific bacteria. J Hazard Mater 151: 339-347.

Chapter 2: Cultivation-independent detection of genes present in soil bacteria

Guo-Chun Ding, Holger Heuer and Kornelia Smalla

Federal Research Centre for Cultivated Plants (JKI), Germany

The development of methods to extract DNA directly from soil samples opened new dimensions of studies on microbial communities in soil. The molecular analysis of soil DNA can provide information on the microbial community composition or on the presence and on the relative abundance of functional genes, e.g., of genes involved in N-cycling, antibiotic resistance genes, degrading genes or plasmid replicon-related sequences. DNA-based analysis of soils led to the discovery of many taxa that are abundant in soils but had not been studied before as they do not easily form colonies on plates. This chapter is exclusively devoted to the detection of genes in DNA directly extracted from soil. Crucial prerequisites of the cultivation-independent detection of genes in soil are appropriate sampling strategies and DNA extraction protocols. Therefore, the subsequent two paragraphs will briefly address these topics before discussing the potentials and limitations of methods for targeted or untargeted detection of genes.

2.1.1 Sampling and Sample Processing

The adequate sampling design depends on the hypotheses to be tested and often pre-experiments might assist in determining the numbers of samples to be analyzed. Sampling of soil requires a careful design of a strategy considering the estimated heterogeneity of the soil system and the expected variability (Van Elsas et al., 2002; Smalla and van Elsas, 2010). To overcome variability due to field heterogeneity, composite samples of soil taken from a plot or a field, often in a number of replicates, were analyzed in many of the studies. The main reason for analyzing composite samples is that the sample should be in a representative manner reporting on the structural and functional diversity of a soil under a particular treatment and on the variation among replicates of the same treatment. Each composite sample may consist of several soil cores taken from a plot in a randomized fashion. Mixing of the soil is usually achieved by sieving the material of the various soil cores through a 2- or 4-mm mesh size. A subsample from each composite sample is then used for the analysis. However, the investigation of composite samples may obscure a meaningful biological variation. Thus, in studies aiming to explore field, plot or micro-scale variability characterization of independent single samples would be required (Sliwinski and Goodman, 2004; Becker et al., 2006).

Nucleic Acid Extraction from Soil Samples

Thirty years after the first paper on the extraction of DNA from soil had been published by Torsvik (1980), obtaining nucleic acids from soil matrices that are suitable for molecular analysis remained a challenge (for review: Van Elsas et al., 2000*)*. None of the protocols for DNA extraction from soil seems suitable for all kinds of soils, in particular for soils originating from contaminated sites. Commercial kits for total DNA extraction from soil facilitated largely the procedure in particular in view of a simplification and miniaturization of the method. Soil DNA extraction kits are often less time consuming and more efficiently remove co-extracted humic acids. Nucleic acids can either be extracted directly from the soil matrix or from a microbial pellet obtained after dislodging microbes associated to soil particles or plant roots and subsequent centrifugation. The first protocol for direct extraction of DNA from soil was published by Ogram et al. (1987) and is based on *in situ* lysis of microbial cells in the presence of the environmental matrix. The second approach, also termed indirect method, was pioneered by Torsvik (1980) and Holben et al. (1988) and involves the extraction of nucleic acids from the microbial fraction. The recovery of the microbial fraction from soils commonly involves repeated homogenization of soil re-suspended in aqueous solution and differential centrifugation steps as originally suggested by Fægri et al. (1977). However, protocols differ considerably with respect to the solutions used to break up soil colloids and dislodge surface-attached cells which adhere to surfaces by various bonding mechanisms such as polymers, electrostatic forces and water bridging, and of different strengths (Bakken and Lindahl, 1995; Van Elsas et al. 2000). This approach might preferably be used for obtaining nucleic acids from soil adhering to roots (rhizosphere), problematic soils (soils with a high content of humic acids or contaminants) or to obtain high molecular DNA for generating fosmid or BAC libraries. Homogenization is usually achieved by shaking soil suspensions with gravel or blending, e.g., with a Stomacher lab blender. Although a complete dislodgment of cells seems to be impossible, it is important that cells bound to soil particles with different degrees of strength are released with similar efficiency. A clear advantage of the indirect approach is that nucleic acids recovered are less contaminated with co-extracted humic acids, DNA of non-bacterial origin or extracellular DNA (unless extracellular DNA is tightly adsorbed to microbial cells). In comparison with the indirect method, the direct DNA extraction approach is

less time consuming and much higher DNA yields are achieved (Van Elsas et al. 2000) that might, however, contain considerable amounts of co-extracted humic acids.

A crucial step for the recovery of representative DNA that mirrors the genomes of all microbes present in a soil sample (Moré et al. 1994; Miller et al., 1999) is the efficient lysis of the bacterial, archaeal and fungal cell walls. Cell lysis can be achieved by mechanical cell disruption and by enzymatic or chemical disintegration of cell walls, or a combination of these methods is often used. The efficiency of the different methods used for cell lysis might not only influence the yield but also the presence of genomic DNA of cells difficult to lyse (resting stage, e.g., dwarf cells or spores). However, obviously the strength of lysis needs to be a trade off as too rigorous lysing methods might shear DNA released from cells that are easy to lyse. Therefore, high molecular weight DNA is an important criterion when evaluating and comparing different protocols because sheared DNA can cause PCR artefacts and is not suitable for direct cloning of large DNA fragments. To obtain large DNA fragments is particularly important for generation of large insert libraries (e.g., fosmid or BAC libraries). In plug lysis of the microbial pellet and pulse field gel electrophoresis is most frequently used to obtain large DNA fragments (Robe et al*.,* 2003). The complete removal of co-extracted humic acids is critical for the subsequent molecular analysis as humic acids were shown to interfere with DNA hybridization, restriction enzyme digestions and PCR amplification (Tebbe and Vahjen, 1993). The DNA yield might vary considerably for different DNA extraction kits used for the same soils. As directly extracted DNA can be of different origin it is recommended not only to estimate the DNA yield on ethidium bromide stained agarose gels or by measuring the ratio 260/280 but also by determining the 16S or 18S rRNA gene copy numbers by quantitative PCR. A recently performed comparison of two frequently used soil DNA extraction kits revealed that the 16S rRNA gene copy numbers were approximately two orders of magnitude different. Differences in DNA yield might be no problem if bacterial, archaeal or fungal ribosomal gene fragments are amplified by PCR. However, if the analysis is targeting less abundant taxa or genes that are carried only by a small proportion of the microbial community then the difference in yield will be critical. Last but not least it should be stressed that strict precautionary measures need to be taken to prevent

contamination of the DNA during the extraction either directly from soil or from the microbial pellet. In particular, when PCR amplification is used to amplify a target gene that occurs less frequently, e.g., antibiotic resistance genes or transgenic DNA in soil, the extraction of DNA, preparation of PCR reactions and analysis of PCR products need to be strictly separated.

2.1.2 Targeted Detection of Genes in Nucleic Acids from Soil

Detection of Genes Providing Information on the Structural Diversity

PCR-based amplification of 16S, 18S rRNA or ITS gene fragments from soil DNA and their subsequent analysis either by cloning and sequencing, direct sequencing or by fingerprinting methods are most frequently used to study the composition of *Bacteria*, *Archaea* or *Fungi* in soil. A considerable advantage is the rapidly growing database of ribosomal rRNA gene sequences which contains presently more than a million good quality 16S rRNA gene sequence entries deposited in Ribosomal Database (http: //rdp.cme.msu.edu, 03/02/2010). A disadvantage of using ribosomal rRNA gene fragments is that bacteria possess different numbers of ribosomal RNA operons that are assumed to reflect different ecological strategies of bacteria (Klappenbach et al., 2000) and that sequence heterogeneity of the different operons might occur (Nübel et al.,1996; von Wintzingerode et al. 1997). Alternative markers such as the gene coding for the σ factor *rpoB* or *gacA* have been suggested since these genes are present only as a single copy and were shown to allow a higher resolution between species (Dahllöf et al., 2000; Costa et al., 2007). However, no matter which gene is targeted, one major limitation that remains is that gene fragments of less common populations are often not represented in clone libraries or fingerprints, especially when primers targeting all bacteria are used for amplification. Bent and Forney (2008) termed this problem “the tragedy of the uncommon” and critically discussed in this context the use of diversity indices. The application of group- or genus-specific primers targeting the 16S rRNA gene can assist in studying less common populations in soil (Heuer et al., 1997, 2001; Costa et al., 2006a; Weinert et al., 2009). In order to be detected by PCR-based methods, the gene copy number of a population should be above 100 or 1000 cells per gram of soil as usually one hundreds or less of the DNA extracted from a gram of soil is used as template

for PCR amplification. Thus the diversity of uncommon populations in soil might become only detectable after soil enrichment steps.

The 16S or 18S rRNA gene or ITS fragments amplified from total community DNA with *Bacteria*, *Archaea* or *Fungi* specific primers were traditionally characterized by cloning and sequencing. Janssen (2006) compared 32 different clone libraries generated from various soils and found that 16S rRNA gene sequences from soil DNA could be affiliated to 32 phyla but that the dominant phyla in the libraries belonged to nine phyla only, i.e., the *Proteobacteria*, *Acidobacteria, Actinobacteria, Verrucomicrobia*, *Bacteroidetes*, *Chloroflexi*, *Planctomycetes*, *Gemmatimonadetes* and *Firmicutes*. Interestingly, the phylum *Firmicutes,* including the classes *Bacilli* and *Clostridia* that were considered to be typical soil bacteria, contributed up to only 2% of the sequences in the clone libraries. However, the composition of the clone libraries, e.g., the relatively low abundance of the *Firmicutes,* might have been strongly affected by DNA extraction biases as mentioned above. Another rather recent methodological development is the use of pyrosequencing to study the diversity of bacterial 16S rRNA gene fragments PCR amplified from soil DNA (Roesch et al., 2007; Fulthorpe et al., 2008; Jones et al., 2009). These studies provided the largest datasets on rRNA gene sequences obtained for soil so far. Furthermore, these data indicated that the number of species per g of soil estimated by Gans et al. (2005) was an overestimate. More importantly, in-depth analysis of soil microbial diversity by sequencing was shown to be no longer impracticable as suggested by Gans et al. (2005). However, despite the enormous amount of data acquired by pyrosequencing, ten thousand sequences are still a very small fraction of $> 10^9$ bacteria living in one gram of soil. Therefore, more attention should be paid to statistical exploration of these datasets. First of all, how well does the dataset reflect the real diversity? Table 5.1.2 demonstrates the relationship among the relative abundance of a species, the total number of sequences and the chance to detect this species. For example, the relative abundance of a species needs to be at least 0.03% to be detected in a sample of 10,000 sequences with a chance of 95%. This corresponds to more than a million cells per gram of soil. However, many soil bacteria will have a lower abundance, at least temporarily or locally. A second important question is how well the dataset from pyrosequencing can represent the community structure. The simulation results in Table 5.1.3 show that there will be a significant error in predicting the relative abundance of rare species. A relative

species abundance of less than 0.01% will be predicted with an error of more than 100%. Obviously, it is still very difficult to reliably predict the community structure also with respect to minorities.

To study spatial and temporal variation of microbial communities in relation to environmental factors, perturbation or experimental treatment, multiple sample analysis is essential (Muyzer and Smalla, 1998; Smalla et al., 2001, 2007; Forney et al., 2004; Kowalchuk et al., 2006). For this purpose approaches based on cloning and sequencing or direct sequencing of 16S rRNA gene fragments amplified from community DNA, are still too costly and labour-intensive, especially for terrestrial habitats of a large microbial diversity. Although a rapid development in the field of high throughput sequencing techniques is witnessed making pyrosequencing more and more affordable and feasible, the use of molecular fingerprinting techniques like denaturing gradient gel electrophoresis (DGGE; TGGE; Muyzer and Smalla, 1998), single strand conformation polymorphism (SSCP; Schwieger and Tebbe, 1998), terminal restriction fragment analysis of 16S ribosomal genes or fragments amplified from total community DNA (T-RFLP; Liu et al., 1997; Osborn et al., 2000) or automated rRNA intergenic spacer analysis (ARISA; Ranjard et al., 2001) remains the state of the art here (Tabel 2.1.1). The molecular fingerprinting techniques based on 16S or 18S rRNA gene or ITS fragments or on alternative marker genes amplified from total community DNA still represent the best compromise between the number of processed samples and the level of information that is obtained. An advantage of microbial community fingerprint analysis is the quick overview of the composition of dominant sequences amplified which reflect the relative abundance of bacterial populations or a specific taxonomic group. When combined with the analysis of sequences of differentiating bands, information on the phylogenetic placement can be rapidly gathered. Recently, Smalla et al. (2007) showed that the analysis of 16S rRNA gene fragments amplified from soil DNA with *Bacteria* specific primers by means of DGGE, SSCP or T-RFLP resulted in comparable results. Although the target regions varied, DGGE, T-RFLP and SSCP analyses all led to similar clustering. In this study little variability in the bacterial community composition of the four replicate composite samples analyzed per soil was observed while significant differences among the different soils were found that mirrored soil physicochemical properties such as pH, clay content or organic C. In a continental-scale study,

including 98 soil samples from North and South America, T-RFLP's of 16S rRNA gene fragments amplified from total community DNA were used to compare bacterial communities across sites (Fierer and Jackson, 2006). The study showed that the composition of soil bacterial communities differed between ecosystems and that most of the differences could be explained by soil pH.

The resolving power of current PCR-based community fingerprinting methods can be considerably increased by using taxon-specific primers. During the PCR reaction with universal primers, a particular target sequence should be present at least at about 0.1% to 1% of the total target sequences to be amplified and appear as a visible band in a fingerprint. To analyze the less common ribotypes and also reduce the complexity of the patterns, nested or semi-nested PCR approaches with taxon-specific primers have been quite successfully used (Heuer et al., 1997; Gomes et al., 2001, 2005; Boon et al., 2002; Salles et al., 2002; Garbeva et al., 2003; Freitag et al., 2005; Costa et al., 2006a, b; Weinert et al., 2009). Different studies that compared DGGE patterns based on16S rRNA gene fragments amplified from DNA or cDNA (following RNA extraction and a reverse transcription step) from soil, reported that DNA- and RNA-based fingerprints of soil bacterial communities are often rather similar (Nicol et al., 2003; Gomes et al., 2005).

Recently, 16S rRNA gene fragments amplified from soil DNA were also analyzed by so-called PhyloChips that enable the parallel detection of more than 8,700 operational taxonomic units (OTU) and the ability to identify individual OTU varying by five orders of magnitude in abundance (Brodie et al., 2006; DeSantis et al., 2007). Typically more than 2,000 OTU can be detected in soil DNA. In the study by DeAngelis et al. (2009) the rhizosphere effect was studied by PhyloChip analysis. OTU responding to the root exudates were detected in 17 of the 44 phyla. PhyloChip analysis of 16S rRNA gene fragments enables a comparison of multiple samples and a determination of taxa responding to biotic or abiotic factors. The great advantage of PhyloChips compared to other fingerprinting methods is that information on the taxonomic affiliation of these responders can be obtained.

However, the molecular fingerprints described above provide no quantitative data on the abundance of the taxa detected. Here the real-time determination of 16S

rRNA gene copy numbers, e.g., of *Bacteria* or *Archaea* or of specific groups is the method of choice. Particular characteristics of the 16S rRNA gene, such as those caused by the presence of different numbers of rRNA operons per genome (from 1 to up to 15), in different species prevent that the copy number determined can be directly translated to cell number. Real time PCR formats that are most commonly used, are SYBR green detection and TaqMan probes. SYBR green binds un-specifically to all double stranded DNA and thus product formation is monitored by an increased fluorescence. In contrast, the TaqMan probe assay is based on the specific binding of the double-labelled probe (quencher- and reporter dye) to a target sequence and the cleavage of the quencher by the Taq polymerase endonuclease activity. Upon cleavage the reporter dye is no longer quenched and the signal can be detected. Again, a sufficient abundance of the target sequence and absence of potential PCR inhibiting substances such as co-extracted humic acids, needs to be considered for a reliable real time PCR detection from soil.

Targeted Detection of Functional Genes

PCR-based detection of functional genes in soil has widely been used in several studies. The presence of different functional genes in soil bacteria has usually been studied by PCR amplification from total DNA and subsequent cloning and sequencing. If sufficient amounts of PCR products can be obtained, the analysis of PCR amplicons by cloning and sequencing or by molecular fingerprints can provide information on the diversity of respective genes, e.g., *pmoA*, (Horz et al., 2001), *dsr* (Schmalenberger et al., 2007), bacterial or archaeal *amoA* (Nicol et al., 2008; Jia and Conrad, 2009), *ndo* genes (Gomes et al., 2007; Flocco et al., 2009), or chitinase genes (Metcalfe et al., 2002).

If the gene targeted is present above a critical detection threshold then its presence can be quantified by real time PCR (Heuer and Smalla, 2007; Heuer et al., 2009; Hai et al., 2009; Schauss et al., 2009). In particular, the development of real time PCR assays enabled studies relating gene abundance to various different biotic or abiotic factors. Again the detection limit is an important issue. Only if gene copies occur with more than 10^5–10^6 copies per gram of soil, amplicons visible in ethidium bromide stained agarose gels are obtained. The detection limit is comparable for real

time PCR. Although SybrGreen detection is widely employed for the detection of functional genes by quantitative PCR, the use of TaqMan probes might increase specificity and sensitivity of the system. In many studies the copy number determined is related to g of soil or even µL of DNA. However, it is strongly recommended that the copy number determined is related to 16S rRNA gene copy numbers to overcome biases caused by uncertainties of the total soil DNA composition. Although comparable amounts of DNA are obtained per gram of soil, the number of 16S rRNA gene copy numbers might vary considerably (Heuer et al., 2009). If the gene targeted is less abundant in the soil population then it is recommendable to increase the sensitivity of their detection by Southern blot hybridization. Even if amplicons are not visible on the agarose gel they can be detected by means of hybridization with digoxigenin or ^{32}P labelled probes. Although the data obtained are at the most semi-quantitative the sensitivity of gene detection is often increased by two orders of magnitude. In addition, by choosing the stringency of the hybridization conditions sequence specificity of the signal can be confirmed. The presence of different functional genes in soil bacteria has been studied by PCR amplification from total DNA and subsequent analysis by Southern blot analysis, e.g., *merA*, *trfA (*Smalla et al., 2006), *nahA*, *nahH, phnAC* (Gomes et al., 2005, 2007).

The spectrum of primers is also a very important issue as the diversity of many genes is still unknown and many novel sequences could be frequently found from environmental samples. The potential for microbial degradation of polycyclic aromatic hydrocarbons (PAH) in the environment is of major interest due to the bioaccumulation and toxicity of these compounds. The diversity of PAH degrading genes in soils and how it is influenced by different soil types or their history and type of pollution is not yet well explored, because the existing primer systems have severe limitations. In Table 5.1.4, published primers that detect PAH ring hydroxylating dioxygenase genes are summarized. Most of the primers target only a rather narrow range of sequences, e.g., *nahAc* or *phnAc* type sequences, or only genes from gram-negative bacteria. One primer system could detect dioxygenase genes of both Gram-positive and Gram-negative bacteria, but the specificity is broader than PAH dioxygenases and the short amplified fragment do not resolve much of the diversity (Ni Chadhain et al., 2006). To overcome these limitations, recently accumulating sequence data need to be used to develop better primer systems to determine the

diversity and abundance of PAH ring hydroxylating dioxygenase genes of soil microbial communities.

2.1.3 Non-targeted Detection of Genes

Functional genes have been detected in soil DNA by the so-called functional arrays. However, the performance of microarray hybridization from complex environments such as soil has to be carefully evaluated and obviously a number of technical challenges need to be overcome before this technique can be fully exploited. A microarray with 100 functional genes was used by Wu et al. (2001) to systematically study the specificity, sensitivity and quantification of microarray hybridization with DNA from complex environmental samples. When environmental DNA is used without prior PCR amplification this seems to be the most difficult challenge. The level of detection is 1,000 to 10,000 fold lower than with PCR amplification (Wu et al., 2001). In order to achieve a comprehensive investigation of functional genes, so-called GeoChip microarrays containing over 24,000 probes covering more than 10,000 genes distributed among more than 150 functional groups involved in nitrogen, carbon, sulphur and phosphorus cycling were recently employed for soil studies (Gentry et al., 2006; Wu et al., 2006; He et al., 2007). The direct hybridizations were, however, not very sensitive, and to increase sensitivity of the approach, a pre-amplification step (e.g., using rolling circle amplification) is now included. Despite of numerous challenges that need to be solved, the GeoChip will be a powerful tool that allows linking functional microbial dynamics to particular ecosystem processes. Probe development, hybridization quality and data evaluation are crucial steps for an appropriate use of DNA microarrays to study the soil microbiota. The GeoChip was applied to study N- and C-cycle genes in DNA extracted from soil taken from different sites in the Antarctic (Yergeau et al., 2007). The N- and C-cycle genes detected differed significantly across sampling locations and vegetation types. However, DNA microarrays cannot generate information on new sequence types and thus only the breadth of known functions can be assessed (Yergeau et al., 2007). A recent study of Suenaga et al. (2009) gives an idea on the severity of this limitation. They created a metagenomic library with DNA extracted from activated sludge and sequenced those clones with extradiol dioxygenase activity. By sequence analysis, they found that most of these clones (36 of 38) contained complete aromatic degradation pathways that shared low similarity to

those found in known cultured bacteria metabolizing aromatic compounds. Seeing that soil bacterial communities can be extremely diverse, it will be a challenge to include all potentially important genetic variants of a particular soil function on a microarray. Furthermore, the abundance of most bacteria may be below the detection limit of microarrays, some of them having the potential to take over important functions when environmental conditions are advantageous for them. To increase the sensitivity, community DNA was frequently subjected to whole genome amplification before microarray analysis.

Another non-targeted approach is cloning large fragments of environmental DNA into fosmid or BAC vectors and subsequent screening of the metagenomic libraries. Metagenomics offers access to functional genes in the soil microbiota and thus to novel bioactive products (Sjöling et al., 2007). This approach has been successfully applied to recover DNA coding for so far unknown enzymes or antibiotics directly from soil DNA (Rondon et al., 2000; Gillespie et al., 2002; Sjöling et al. 2007). However, despite the vast potential of this approach, several methodological challenges remain to be solved. To obtain sequence information of particular uncultured taxa which are assumed to be abundant in the soil sample, in a first step specific primers are used to identify the clones carrying the respective genes, either based on 16S rRNA or even on functional genes. Using this approach, genomic fragments of uncultured *Acidobacteria* or *Archaea* could be obtained and subjected to sequence-based gene detection (Ochsenreiter et al., 2003; Treusch et al., 2005; Van Elsas et al., 2008a, b). Sequencing of a large fosmid library from a grassland soil revealed that a fragment with an rRNA gene, which showed the highest similarity to group 1.1b of the crenarchaeota, carried genes encoding copper nitrite reductase and two subunits of ammonia monooxygenase or particulate methane monooxygenase (Treusch et al., 2005). The discovery of ammonia oxidizing *Archaea* showed the potential of soil metagenomics studies to elucidate the role of uncultured organisms in, e.g., soil nutrient cycles (Leininger et al., 2006). Considering that the majority of soil bacteria occur at relatively low abundance but might have an important ecological role, the development of approaches such as targeted access to low-abundance bacteria are of considerable interest. Recently Hjort et al. (2010) analyzed chitinase genes in DNA directly extracted from a phytopathogen-suppressive soil, in a metagenomic library constructed from microbial cells extracted

from soil, and in genomic DNA from bacterial isolates with antifungal and chitinase activities. Although T-RFLP of chitinase genes revealed differences in amplified chitinase genes from the metagenomic library and the directly extracted DNA, approximately 40% of the identified chitinase T-RFs were found in both sources. All of the chitinase T-RFs from the isolates were matched to T-RFs in the directly extracted DNA and the metagenomic library. The authors demonstrated an impressive agreement between three very different screening techniques all of which pointed towards specific *Streptomyces* species that could play a role in suppression of fungi by chitinase production in soil. However, some clusters of chitinase genes were represented depending on the approach used. Therefore, Hjort et al. (2010) concluded that a combination of molecular approaches increases the information obtained and the reliability of the data.

2.1.4 Cultivation-independent Detection of Genes Carried by Mobile Genetic Elements

The importance of horizontal gene exchange for short-term bacterial adaptability and for successful colonization of new ecological niches has only recently been appreciated (Heuer et al., 2008). Nowadays mobile genetic elements (MGE) are recognized as an important and essential component that promotes bacterial diversity. The PCR-based detection of MGE-specific sequences in community DNA was first used by Götz et al. (1996). Primers targeting replicon-specific sequences were designed on the basis of sequenced broad host range plasmids or integrons. In combination with Southern blot hybridization a specific and sensitive monitoring of large numbers of environmental samples became possible (Smalla et al., 2000; 2006; Heuer et al., 2002, 2009). While the targeted detection of MGE by endpoint or real time PCR requires sequence information, capturing MGE directly from soil microbial pellets and subsequent sequencing enables a non-targeted discovery and detection of genes. Direct capturing of MGE by means of so-called exogenous plasmid isolation techniques (Bale et al., 1988; Hill et al., 1992) that has been widely used to capture MGE conferring selectable traits such as mercury or antibiotic resistance have been acquired from a wide range of soil or sediments in gram-negative recipients functioning as a genetic sink (Heuer et al., 2002, 2009; Smalla et al., 2006; Binh et al., 2008). Capturing of degradative genes resident on MGE has

been demonstrated as well (Top et al., 1995). Increased transfer frequencies have often been observed when the soil environmental sample was previously exposed to pollutants. The PCR-based detection of MGE has been used for monitoring the abundance of MGE in soils (Götz et al., 1996; Smalla et al., 2000; Heuer et al., 2009), and “hot spots” with high abundance of MGE could be identified.

2.1.5 Conclusion and Outlook

The tools for cultivation-independent detection of genes in soil bacteria rapidly advanced over the last few years. The technological developments briefly described in this chapter brought us the opportunity to study the enormous complexity of soil microbial communities in more comprehensive and complete terms. The use of molecular tools has dramatically changed our view of the microbes residing in soil. Most importantly, it became possible not only to study who is there but — although still a methodological challenge — also what are they doing. The advances of soil metagenomic tools will continue to improve our understanding of soil microbes. To better understand the enormous microbial diversity and the biotic and abiotic factors shaping their composition and functions it is certainly needed to study the soil microbiota at different scale and also to develop approaches to target the uncommon or rare populations in soil. The rapidly advancing tools will assist to successfully uncover more secrets of microbial life in soil.

References

Bakken, L.R., and V. Lindahl. 1995. Recovery of bacterial cells from soil. P. 13-27. In J.T. Trevors, and J.D. van Elsas (ed.) Nucleic acids in the environment. Springer. Heidelberg, Germany

Baldwin, B.R., C.H. Nakatsu, and L. Nies. 2003. Detection and enumeration of aromatic oxygenase genes by multiplex and real-time PCR. Appl. Environ. Microbiol. 69:3350-3358.

Bale, M.J., M.J. Day, and J.C. Fry. 1988. Novel method for studying plasmid transfer in undisturbed river epilithon. Appl. Environ. Microbiol. 54:2756-2758.

Becker, J.M., T. Parkin, C.H. Nakatsu, J.D. Wilbur, and A. Konopka. 2006. Bacterial activity, community structure, and 48luorine48s-scale spatial heterogeneity in contaminated soil. Microb. Ecol. 51: 220-231.

Bent, S.J., and L.J. Forney. 2008. The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. ISME J. 2:689-695.

Binh, C.T.T., H. Heuer, M. Kaupenjohann, and K. Smalla. 2008. Piggery manure used for soil fertilization is a reservoir for transferable antibiotic resistance plasmids. FEMS Microbiol. Ecol. 66:25-37.

Bordenave, S., M. Goni-Urriza, C. Vilette, S. Blanchard, P. Caumette, and R. Duran. 2008. Diversity of ring-hydroxylating dioxygenases in pristine and oil contaminated microbial mats at genomic and transcriptomic levels. Environ. Microbiol. 10:3201-3211.

Boon, N., W. De Windt, W. Verstraate, and E.M. Top. 2002. Evaluation of nested PCR-DGGE (denaturing gradient gel electrophoresis) with group-specific 16S rRNA primers for the analysis of bacterial communities from different wastewater treatment plants. FEMS Microbiol. Ecol. 39:101-112.

Brodie, E.L., T.Z. DeSantis, D.C. Joyner, S.M. Baek, J.T. Larsen, G.L. Andersen, T.C. Hazen, P.M. Richardson, D.J. Herman, T.K. Tokunaga, J.M.M. Wan, and M.K. Firestone. 2006. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Appl. Environ. Microbiol. 72:6288-6298.

Cebron, A., M.P. Norini, T. Beguiristain, and C. Leyval. 2008. Real-Time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHDa) genes from Gram positive and Gram negative bacteria in soil and sediment samples. J. Microbiol. Methods 73: 148-159.

Costa, R., M. Götz, N. Mrotzek, J. Lottmann, G. Berg, and K. Smalla. 2006a. Effect of site and plant species on rhizosphere community structure as revealed by molecular analysis of different microbial guilds. FEMS Microbiol. Ecol. 56:236-249.

Costa, R., J.F. Salles, G. Berg, and K. Smalla. 2006b. Cultivation-independent analysis of Pseudomonas species in soil and in the rhizosphere of field-grown *Verticillium 48luori* host plants. Environ. Microbiol. 8:2136-2149.

Costa, R., N.C.M. Gomes, E. Krögerrecklenfort, K. Opelt, G. Berg, and K. Smalla. 2007. Pseudomonas community structure and antagonistic potential in the rhizosphere: insights gained by combining phylogenetic and functional gene-based analyses. Environ. Microbiol. 9:2260-2273.

Dahllöf, I., H. Baillie, and S. Kjelleberg. 2000. rpoB-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. Appl. Environ. Microbiol. 66:3376-3380.

DeAngelis, K.M., E.L. Brodie, T.Z. DeSantis, G.L. Andersen, S.E. Lindow, and M.K. Firestone. 2009. Selective progressive response of soil microbial community to wild oat roots. ISME J. 3:168-178.

Debruyn, J.M., C.S. Chewning, and G.S. Sayler. 2007 Comparative quantitative prevalence of Mycobacteria and functionally abundant *nidA, nahAc,* and *nagAc* dioxygenase genes in coal tar contaminated sediments. Environ. Sci. Technol. 41:5426-5432.

DeSantis, T.Z., E.L. Brodie, J.P. Moberg, I.X. Zubieta, Y.M. Piceno, and G.L. Andersen. 2007. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microb. Ecol. 53:371-383.

Dionisi, H.M., C.S. Chewning, K.H. Morgan, F.M. Menn, J.P. Easter, and G.S. Sayler. 2004. Abundance of dioxygenase genes similar to Ralstonia sp. Strain U2 nagAc is correlated with naphthalene concentrations in coal tar-contaminated freshwater sediments. Appl. Environ. Microbiol. 70:3988-3995.

Fægri, A., V.L. Torsvik, and J. Goksøyr. 1977. Bacterial and fungal activities in soil: separation of bacteria and fungi by a rapid fractionated centrifugation technique. Soil Biol. Biochem. 9:105-112.

Fierer, N., and R.B. Jackson. 2006. The diversity and biogeography of soil bacterial communities. PNAS 103:626-631.

Flocco, C.G., N.C.M. Gomes, W.M. Cormack, and K. Smalla. 2009. Occurrence and diversity of naphthalene dioxygenase genes in soil microbial communities from the Maritime Antarctic. Environ. Microbiol. 11:700-714.

Forney, L.J., X. Zhou, and C.J. Brown. 2004. Molecular microbial ecology: land of the one-eyed king. Curr. Opin. Microbiol. 7:210-220.

Freitag, T.E., L. Chang, C.D. Clegg, and J.I. Prosser. 2005. Influence of inorganic nitrogen management regime on the diversity of nitrite-oxidizing bacteria in agricultural grassland soils. Appl. Environ. Microbiol. 71:8323-8334.

Fulthorpe, R.R., L.F.W. Roesch, A. Riva, and E.W. Triplett. 2008. Distantly sampled soils carry few species in common. ISME J. 2:901-910.

Gans, J., M. Wolinsky, and J. Dunbar. 2005. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 309:1387-1390.

Garbeva, P., J.A. Van Veen, and J.D. Van Elsas. 2003. Predominant *Bacillus* spp. In agricultural soil under different management regimes detected via PCR-DGGE. Microb. Ecol. 45:302-316.

Gentry, T.J., G.S. Wickham, C.W. Schadt, Z. He, and J. Zhou. 2006. Microarray applications in microbial ecology research. Microb. Ecol. 52: 159-175.

Gillespie, D.E., S.F. Brady, A.D. Bettermann, N.P. Cianciotto, M.R. Liles, M.R. Rondon, J. Clardy, R.M. Goodman, and J. Handelsman. 2002. Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. Appl. Environ. Microbiol. 68:4301-4306.

Götz, A., R. Pukall, E. Smit, E. Tietze, R. Prager, H. Tschäpe, J.D. van Elsas, and K. Smalla. 1996. Detection and characterization of broad-host range plasmids in environmental bacteria by PCR. Appl. Environ. Microbiol. 62: 2621-2628.

Gomes, N.C.M., H. Heuer, J. Schönfeld, R. Costa, L. Hagler-Mendonça, and K. Smalla. 2001. Bacterial diversity of the rhizosphere of maize (Zea mays) grown in tropical soil studied by temperature gradient gel electrophoresis. Plant Soil 232: 167-180.

Gomes, N.C.M., I.A. Kosheleva, W.-R. Abraham, and K. Smalla. 2005. Effects of the 50luorine50s strain Pseudomonas putida KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. FEMS Microbiol. Ecol. 54: 21-33.

Gomes, N.C.M., L.R. Borges, R.Paranhos, F.N. Pinto, E. Krögerrecklenfort, L.C.S. Mendonça-Hagler, and K. Smalla. 2007. Diversity of *ndo* genes in mangrove sediments exposed to different sources of polycyclic aromatic hydrocarbon pollution. Appl. Environ. Microbiol. 73:7392-7399.

Hai, B., N.H. Diallo, S. Sall, F. Haesler, K. Schauss, M. Bonzi, K. Assigbetse, J.-L. Chotte, J.C. Munch, and M. Schloter. 2009. Quantification of key genes steering the microbial nitrogen cycle in the rhizosphere of 2 sorghum cultivars in tropical agro-ecosystems. Appl. Environ. Microbiol. 75:4993-5000.

He, Z., T.J. Gentry, C.W. Schadt, L. Wu, J. Liebich, S.C. Chong, Z. Huang, W. Wu, B. Gu, P. Jardine, C. Criddle, and J. Zhou. 2007. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J. 1:67-77.

Heuer, H., and K. Smalla. 2007. Manure and sulfadiazine synergistically increased bacterial antibiotic resistance in soil over at least two months. Environ. Microbiol. 9:657-666.

Heuer, H., Z. Abdo, and K. Smalla. 2008. Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. FEMS Microbiol. Ecol. 65:361-371.

Heuer, H., M. Krsek, P. Baker, K. Smalla, and E.M.H. Wellington. 1997. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl. Environ. Microbiol. 63:3233-3241.

Heuer, H., G. Wieland, J. Schönfeld, A. Schönwälder, N.C.M. Gomes, and K. Smalla. 2001. Bacterial community profiling using DGGE or TGGE analysis. P. 177-190. In P. Rouchelle (ed.) Environmental molecular microbiology: Protocols and applications. Horizon Scientific Press. Wymondham, UK.

Heuer, H., E. Krögerrecklenfort, S. Egan, L.S. van Overbeek, G. Guillaume, T.L. Nikolakopoulou, E.M.H. Wellington, J.D. van Elsas, J.-M. Collard, A.D. Karagouni, and K. Smalla. 2002. Gentamicin resistance genes in environmental bacteria: Prevalence and transfer. FEMS Microbiol. Ecol. 42:289-302.

Heuer, H., C. Kopmann, C.T.T. Binh, E.M. Top, and K. Smalla. 2009. Spreading antibiotic resistance through spread manure: characteristics of a novel plasmid type with low %G+C content. Environ. Microbiol. 11: 937-949.

Hill, K.E., A.J. Weightman, and J.C. Fry. 1992. Isolation and screening of plasmids from the epilithon which mobilize recombinant plasmid pD10. Appl. Environ. Microbiol. 58:1292-1300.

Hjort, K., M. Bergström, M.F. Adesina, J.K. Jansson, K. Smalla, and S. Sjöling. 2010. Chitinase genes revealed and compared in bacterial isolates, DNA extracts and a metagenomic library from a phytopathogen suppressive soil. FEMS Microbiol. Ecol. 71:197-207.

Holben, W.E., J.K. Jansson, B.K. Chelm, and J.M. Tiedje. 1988. DNA probe method for the detection of specific microorganisms in the soil bacterial community. Appl Environ. Microbiol. 54:703-711.

Horz, H.-P., M. Tchawa Yimaga, and W. Liesack. 2001. Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of pmoA, mmoX, mxaF and 16S rRNA and ribosomal DNA, including pmoA-based terminal restriction fragment length polymorphism profiling. Appl. Environ. Microbiol. 67:4177-4185.

Janssen, P.H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. Appl. Environ. Microbiol. 72:1719-1728.

Jia, Z., and R. Conrad. 2009. Bacteria rather than Archaea dominate microbial ammonia oxidation in an agricultural soil. Environ. Microbiol. 11:1658-1671.

Jones, R.T., M.S. Robeson, C.L. Lauber, M. Hamady, R. Knight, and N. Fierer. 2009. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. ISME J. 3:442-453.

Klappenbach, J.A., J.M. Dunbar, and T.M. Schmidt. 2000. rRNA operon copy number reflects ecological strategies of bacteria. Appl. Environ. Microbiol. 66:1328-1333.

Kowalchuk, G.A., W.H.G. Hol, and J.A. van Veen. 2006. Rhizosphere fungal communities are influenced by Senecio jacobaea pyrrolizidine alkaloid content and composition. Soil Biol. Biochem. 38:2852-2859.

Laurie, A.D., and G. Lloyd-Jones. 2000. Quantification of phnAc and nahAc in contaminated New Zealand soils by competitive PCR. Appl. Environ. Microbiol. 66:1814-1817.

Leininger, S., T. Urich, M. Schloter, L. Schwark, J. Qi, G.W. Nicol, J.I. Prosser, S.C. Schuster, and C. Schleper. 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442:806-809.

Liu, W.-T., T.L. Marsh, H. Cheng, and L.J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. 63:4516-4522.

Lozada, M., J.P. Riva Mercadal, L.D. Guerrero, W.D. Di Marzio, M.A. Ferrero, and H.M. Dionisi. 2008. Novel aromatic ring-hydroxylating dioxygenase genes from coastal marine sediments of Patagonia. BMC Microbiol. 8:50.

Metcalfe, A.C., M. Krsek, G.W. Gooday, J.I. Prosser, and E.M.H. Wellington. 2002. Molecular analysis of a bacterial chitinolytic community in an upland pasture. Appl. Environ. Microbiol. 68:5042-5050.

Miller, D.N., J.E. Bryant, E.L. Madsen, and W.C. Ghiorse. 1999. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Appl. Environ. Microbiol. 65:4715-4724.

Moré, M.I., J.B. Herrick, M.C. Silva, W.C. Ghiorse, and E.L. Madsen. 1994. Quantitative cell lysis of indigenous microorganisms and rapid extraction of microbial DNA from sediment. Appl. Environ. Microbiol. 60:1572-1580.

Muyzer, G., and K. Smalla. 1998. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. A. Van Leeuwenhoek 73:127-141.

Ni Chadhain, S.M., R.S. Norman, K.V. Pesce, J.J. Kukor, and G.J. Zylstra. 2006. Microbial dioxygenase gene population shifts during polycyclic aromatic hydrocarbon biodegradation. Appl. Environ. Microbiol. 72:4078-4087.

Nicol, G.W., L.A. Glover, and J.I. Prosser. 2003. Spatial analysis of archaeal community structure in grassland soil. Appl. Environ. Microbiol. 69:7420-7429.

Nicol, G.W., S. Leininger, and C. Schleper. 2008. The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. Environ. Microbiol. 10:2966-2978.

Nübel, U., B. Engelen, A. Felske, J. Snaidr, A. Wiesenhuber, R.I. Amann, W. Ludwig, and H. Backhaus. 1996. Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. J. Bacteriol. 178: 5636-5643.

Ochsenreiter, T., D. Selezi, A. Quaiser, L. Bonch-Osmolovskaya, and C. Schleper. 2003. Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. Environ. Microbiol. 5: 787-797.

Ogram, A., G.S. Sayler, and T.J. Barkay. 1987. DNA extraction and purification from sediments. J. Microbiol. Methods 7:57-66.

Osborn, A.M., E.R.B. Moore, and K.N. Timmis. 2000. An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ. Microbiol. 2:39-50.

Ranjard, L., F. Poly, J.C. Lata, C. Mougel, J. Thioulouse, and S. Nazaret. 2001. Characterization of bacterial and fungal soil communities by automated ribosomal intergenic spacer analysis fingerprints: biological and methodological variability. Appl. Environ. Microbiol. 67:4479-4487.

Robe, P., R. Nalin, C. Capellano, T.A. Vogel, and P. Simonet. 2003. Extraction of DNA from soil. Eur. J. Soil Biol. 39: 183-190.

Roesch, L.F.W., R.R. Fulthorpe, A. Riva, G. Casella, A.K.M. Hadwin, A.D. Kent, S.H. Daroub, F.A.O. Camargo, W.G. Farmerie, and E.W. Triplett. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J. 1: 283-290.

Rondon, M.R., P.R. August, A.D. Bettermann, S.F. Brady, T.H. Grossman, M.R. Liles, K.A. Loiacono, B.A. Lynch, I.A. MacNeil, M.S. Osburne, J. Clardy, J. Handelsman, and R.M. Goodman. 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl. Environ. Microbiol. 66:2541-2547.

Salles, J.F., F.A. de Souza, and J.D. van Elsas. 2002. Molecular method to assess the diversity of Burkholderia species in environmental samples. Appl. Environ. Microbiol. 68:1595-1603.

Schauss, K., A. Focks, S. Leininger, A. Kotzerke, H. Heuer, S. Thiele-Bruhn, S. Sharma, B.-M. Wilke, M. Matthies, K. Smalla, J.C. Munch, W. Amelung, M. Kaupenjohann, M. Schloter, and C. Schleper. 2009. Dynamics and functional relevance of ammonia-oxidizing archaea in two agricultural soils. Environ. Microbiol. 11:446-456.

Schmalenberger, A., H.L. Drake, and K. Küsel. 2007. High unique diversity of 53luorine-reducing prokaryotes characterized in a depth gradient in an acidic fen. Environ. Microbiol. 9:1317-1328.

Schwieger, F., and C.C. Tebbe. 1998. A new approach to utilize PCR-single strand-conformation polymorphism for 16S rRNA gene-based microbial community analysis. Appl. Environ. Microbiol. 64:4870-4876.

Sjöling, S., W. Stafford, and D.A. Cowan. 2007. Soil metagenomics: Exploring and exploiting the soil microbial gene pool. In p.409-434. J.D. van Elsas, J. Jansson, and J.T. Trevors (ed.) Modern soil microbiology. 2nd Ed. CRC Press, Taylor and Francis Group. Boca Raton, FL.

Sliwinski, M.K., and R.M. Goodman. 2004. Spatial heterogeneity of Crenarchaeal assemblages within mesophilic soil ecosystems as revealed by PCR-Single-Stranded Conformation Polymorphism profiling. Appl. Environ. Microbiol. 70:1811-1820.

Smalla, K., and J.D. van Elsas. 2010. The soil environment. P. 111-130. In W.T. Liu, and J.K. Janson (ed.) Environmental molecular microbiology. Caister Academic Press. Norfolk, UK.

Smalla, K., H. Heuer, A. Götz, D. Niemeyer, E. Krögerrecklenfort, and E. Tietze. 2000. Exogenous isolation of antibiotic resistance plasmids from piggery manure slurries reveals a high prevalence and diversity of IncQ-like plasmids. Appl. Environ. Microbiol. 66:4854-4862.

Smalla, K., G. Wieland, A. Buchner, A. Zock, J. Parzy, S. Kaiser, N. Roskot, H. Heuer, and G. Berg. 2001. Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. Appl. Environ. Microbiol. 67:4742-4751.

Smalla, K., A.S. Haines, K. Jones, E. Krögerrecklenfort, H. Heuer, M. Schloter, and C.M. Thomas. 2006. Increased abundance of IncP-1β plasmids and mercury resistance genes in mercury-polluted river sediments: First discovery of IncP-1β plasmids with a complex mer transposon as the soil accessory element. Appl. Environ. Microbiol. 72:7253-7259.

Smalla, K., M. Oros-Sichler, A. Milling, H. Heuer, S. Baumgarte, R. Becker, G. Neuber, S. Kropf, A. Ulrich, and C.C. Tebbe. 2007. Bacterial diversity of soils assessed by DGGE, T-RFLP and SSCP fingerprints of PCR-amplified 16S rRNA gene fragments: do the different methods provide similar results? J. Microbiol. Methods 69:470-479.

Suenaga, H., Y. Koyama, M. Miyakoshi, R. Miyazaki, H. Yano, M. Sota, Y. Ohtsubo, M. Tsuda, and K. Miyazaki. 2009. Novel organization of aromatic degradation pathway genes in a microbial community as revealed by metagenomic analysis. ISME J. 3:1335-1348.

Tebbe, C.C., and W. Vahjen. 1993. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. Appl. Environ. Microbiol. 59:2657-2665.

Top, E. M., W.E. Holben, and L.J. Forney. 1995. Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. 61:1691-1698.

Torsvik, V. 1980. Isolation of bacterial DNA from soil. Soil Biol. Biochem. 12:15-21.

Treusch, A.H., S. Leininger, A. Kletzin, S.C. Schuster, H.-P. Klenk, and C. Schleper. 2005. Novel genes for nitrite reductase and *amoA*-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. Environ. Microbiol. 7:1985-1995.

Van Elsas, J.D., K. Smalla, and C.C. Tebbe. 2000. Extraction and analysis of microbial community nucleic acids from environmental matrices. P. 29-51. In J.K. Jansson, J.D. van Elsas, and M.J. Bailey (ed.) Tracking genetically-engineered microorganisms. Eurekah. Austin, TX.

Van Elsas, J.D., K. Smalla, A.K. Lilley, and M.J. Bailey. 2002. Methods for sampling soil microbes. P. 505-526. In C.J. Hurst, R.L. Crawford, G.R. Knudsen, M.J. McInerney, and L.D. Stetzenbach (ed.) Manual of environmental microbiology, 2nd Ed. ASM Press. Washington, DC.

Van Elsas, J.D., A.J. Speksnijder, and L.S. van Overbeek. 2008a. A protocol for the metagenomic analysis of suppressive soil. J. Microbiol Methods 75:515-522.

Van Elsas, J.D., R. Costa, J.K. Jansson, S. Sjöling, M.J. Bailey, R. Nalin, T.M. Vogel, and L.S. van Overbeek. 2008b. The metagenomics of disease-suppressive soils – Experiences from the METACONTROL project. Trends Biotechnol. 26:591-601.

Von Wintzingerode, F., U.B. Göbel, and E. Stackebrandt. 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev. 21:213-229.

Weinert, N., R. Meincke, C. Gottwald, H. Heuer, N.C. Gomes, M. Schloter, G. Berg, and K. Smalla. 2009. Rhizosphere communities of genetically modified zeaxanthin-accumulating potato plants and their parent cultivar differ less than those of different potato cultivars. Appl. Environ. Microbiol. 75:3859-3865.

Wilson, M.S., C. Bakermans, and E.L. Madsen. 1999. In situ, real-time catabolic gene expression: extraction and characterization of naphthalene dioxygenase mRNA transcripts from groundwater. Appl. Environ. Microbiol. 65:80-87.

Wu, L., D.K. Thompson, G. Li, R.A. Hurt, J.M. Tiedje, and J. Zhou. 2001. Development and evaluation of functional gene arrays for detection of selected genes in the environment. Appl. Environ. Microbiol. 67:5780-5790.

Wu, L.Y., X. Liu, C.W. Schadt, and J.Z. Zhou. 2006. Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl. Environ. Microbiol. 72:4931-4941.

Yergeau, E., S. Kang, Z. He, J. Zhou, and G.A. Kowalchuk. 2007. Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. ISME J. 1:163-179.

Zhou, H.W., C.L. Guo, Y.S. Wong, and N.F. Tam. 2006. Genetic diversity of dioxygenase genes in polycyclic aromatic hydrocarbon-degrading bacteria isolated from mangrove sediments. FEMS Microbiol. Lett. 262:148-157.

Table 2.1.1 Molecular techniques used to compare and analyse microbial communities from soil.

Technique	Principles	Reference
DGGE, TGGE	Denaturing gradient gel electrophoresis: PCR-amplified fragments of 16S/18S rRNA, RpoB, GacA or other genes are separated based on their melting in a denaturing gradient formed by chemicals (DGGE) or temperature (TGGE)	Muyzer and Smalla (1998)
SSCP	Single strand conformation polymorphism: digestion of one DNA-strand of PCR-amplified fragments, and electrophoretic separation based on the sequence determined conformation of the remaining single strand	Schwieger and Tebbe (1998)
T-RFLP	Terminal restriction fragment length polymorphism: PCR-amplification of gene fragments using fluorescently labelled primers, cutting off the labelled terminal fragments using restriction enzymes, and separation on a sequencer based on the size of the terminal fragments	Liu et al. (1997)
ARISA	Automated rRNA intergenic spacer analysis: PCR-amplification of intergenic spacers of ribosomal operons using fluorescently labelled primers targeting rRNA genes, and separation on a sequencer based on the size of the fragments	Ranjard et al. (2001)
Pyro-sequencing of rRNA genes	Massive sequencing of PCR-amplified 300-500 base fragments by synthesizing a complementary strand and detection of pyrophosphate release on nucleotide incorporation.	Roesch et al. *(*2007)

Table 2.1.2 Relative abundance of a species that is needed to detect its sequence within a total number of sequences at a specified likelihood, based on a Poisson distribution.

Likelihood of detection	Minimal relative abundance			
	Total number of sequences			
	10	100	1000	10000
0.1	1.0%	0.11%	0.01%	0.001%
0.3	3.5%	0.36%	0.04%	0.004%
0.5	6.7%	0.69%	0.07%	0.007%
0.7	11.3%	1.20%	0.12%	0.012%
0.9	20.6%	2.28%	0.23%	0.023%
0.95	25.9%	2.95%	0.30%	0.030%
0.99	36.9%	4.50%	0.46%	0.046%

Table 2.1.3 Uncertainty to predict the community structure from pyrosequencing data (coefficient of variance from 1,000 simulations to sample 9 species of a particular relative abundance).

Relative abundance	Error of species abundance prediction			
	Total number of sequences			
	1000	2000	10000	20000
10%	9.7%	6.9%	3.1%	2.2%
1%	30.9%	21.5%	9.5%	7.0%
0.1%	95.4%	66.8%	30.8%	21.9%
0.01%	262.1%	209.6%	101.7%	71.7%

Table 2.1.4 PCR systems to study the diversity and abundance of PAH dioxygenase (PAH-RHDa) genes in soil.

Primers	Target genes	Amplicon size	Sample sources	Sample type and analysis	Reference
NAPH-1F / -1R NAPH-2F / -2RGC	*nahAc, phnAc, nagAc, ndo* (gram negative bacteria)	896 bp 740 bp	PAH contaminated mangrove sediments (Brazil) or soils (Maritime Antarctic)	Total community DNA, DGGE and clone library	Gomes et al. (2007)
FRT5A / FRT3B	PAH-RHDa genes of gram negative bacteria	437 bp	long-term and short-term oil contaminated microbial mats	Total community DNA, clone library	Bordenave et al. (2008)
FRT6A / FRT4B	PAH-RHDa genes of gram negative bacteria	491 bp			
Nah-for Nah-rev1 / -rev2	PAH-RHDa genes *nahAc, phnAc,* gram negative bacteria	937 bp 317 bp	PAH enrichment of mangrove sediments, China	PAH degrading isolates, PCR and sequencing	Zhou et al. (2006)
Nid-for Nid-rev1 / -rev2	PAH-RHDa genes of gram positive bacteria	~600 bp ~310 bp			
PAH-RHDα GN F610a+ R16a	PAH-RHDa genes of gram negative bacteria	360 bp	Five kinds of PAH contaminated soil with different texture, France	Total community DNA, Clone library and real time PCR	Cebron et al. (2008)
PAH-RHDα GP F 641b + R 933	PAH-RHDa genes of gram positive bacteria	292 bp			
Rieske_f Rieske_r	(1,2,4-6)57luorin centre of PAH-RHDa genes	78 bp	PAH enrichment from contaminated soils, USA	Total DNA from enriched cell pellets, clones	Ni Chadhain et al. (2006)
Cyc372F Cyc854R	*phnA1*-like genes	479-482 bp	PAH contaminated intertidal sediments	Total community DNA, clone library	Lozada et al. (2008)
nahAcfor / -rev	*nahAc*-like genes	1009 bp	PAH contaminated soil microcosms,New Zealand	Total community DNA, clones and real time PCR	Laurie et al. (2000)
P8073 / P9047	*phnAc*-like genes	974 bp			
Ac114F Ac596R	*nahAc* –like genes	482 bp	Groundwater, possibly contaminated	RNA from extracted cells, real time PCR and dot blot	Wilson et al. (1999)
NAH-F NAH-R	*nahAc*-like genes	377 bp	Isolates	Genomic DNA, real time PCR and hybridization	Baldwin et al. (2003)
nidA-F / -R nahAc-F / -R	*nidA* *nahAc*	~100 bp	Coal tar contaminated sediments	Total community DNA, real time PCR (TaqMan probes)	Debruyn et al. (2007)
nagAc-like-F nagAc-like-R	*nagAc* like genes	107 bp	Coal tar contaminated freshwater sediments	Total community DNA, real time PCR (TaqMan probe)	Dionisi et al. (2004)

Chapter 3: Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes using a novel PCR detection system

Guo-Chun Ding,[1] Holger Heuer,[1] Sebastian Zühlke,[2] Michael Spiteller,[2] Geertje Johanna Pronk,[3] Katja Heister,[3] Ingrid Kögel-Knabner,[3] and Kornelia Smalla[1]*

Julius Kühn-Institut (JKI)—Federal Research Centre for Cultivated Plants, Messeweg 11-12, 38104 Braunschweig, Germany[1]*;*

Institute of Environmental Research, TU Dortmund, 44221 Dortmund, and Lehrstuhl für Bodenkunde, Technische Universität München, 85350 Freising-Weihenstephan, Germany[3]

Abstract:

A novel PCR primer system that targets a wide range of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase (PAH-RHDα) genes of both Gram-positive and Gram-negative bacteria was developed and used to study their abundance and diversity in two different soils in response to phenanthrene spiking. The specificities and target ranges of the primers predicted *in silico* were confirmed experimentally by cloning and sequencing of PAH-RHDα gene amplicons from soil DNA. Cloning and sequencing showed the dominance of *phnAc* genes in the contaminated Luvisol. In contrast, high diversity of PAH-RHDα genes of Gram-positive and Gram-negative bacteria was observed in the phenanthrene-spiked Cambisol. Quantitative real-time PCR based on the same primers revealed that 63 days after phenanthrene spiking, PAH-RHDα genes were one order of magnitude more abundant in the Luvisol than in the Cambisol, while they were not detected in both control soils. In conclusion, sequence analysis of the amplicons obtained confirmed the specificity of the novel primer system and revealed a soil type-dependent response of PAH-RHDα gene-carrying soil bacteria to phenanthrene spiking.

Introduction

Polycyclic aromatic hydrocarbons (PAHs) are hydrophobic compounds composed of two or more fused aromatic rings. Although PAHs are ubiquitous in the environment (from natural oil seeps, bush fires, and plant derivatives), anthropogenic activities, such as disposal of coal-processing waste, mining accidents, petroleum wastes, and vehicle exhaust, have drastically increased their occurrence in the environment. The fate of PAHs in soil is of great interest due to their potential for bioaccumulation, persistence, transport, and toxicity. Microbe-driven aerobic degradation of PAHs is well documented (15–17). The diversity of PAH-degrading genes in soils is assumed to be huge, but the extent of diversity and how it is influenced by different soil types or their history and type of pollution are not yet fully explored. Knowledge of the genes coding for dioxygenase enzymes that catalyze the primary step of PAH degradation by incorporating molecular oxygen into the aromatic nucleus is an essential prerequisite for unraveling the contributions of microbial population networks to transformation, assimilation, and degradation of organic chemicals in soil. Recently, the complete genomes of several PAH-degrading bacteria became available and allowed new insights into degradative pathways (6, 18, 36). Organic pollutants also serve as nutrients for those microbes that have the appropriate genetic makeup to utilize them, resulting in their increased metabolic activity and abundance (4, 14). In the last decade, impressive progress was seen in techniques that allow cultivation-independent analysis of microbial communities and thus overcome the most severe limitations in studying microbial communities in natural habitats, namely, that only a rather small portion of microbes are accessible to standard cultivation conditions (1, 29). For more than a decade, cultivation-independent approaches have also been employed to unravel the responses of microbial communities in soils and sediments to PAH pollution. In all these studies, PCR amplification of PAH-degrading gene fragments from nucleic acids directly extracted from environmental samples was used to explore the abundance and diversity of PAH ring-hydroxylating dioxygenase (PAH-RHDα) genes (4, 8, 9, 13, 14, 22, 34, 37). Despite the known biases of PCR amplification from mixed templates, these techniques allow highly sensitive and specific detection even from minute amounts of nucleic acids. In order to select suitable primer systems, previously published primer systems were analyzed for their ranges of target sequences. The

existing primer systems were found to have limitations, as they often target only a rather narrow range of sequences, e.g., *nahAc*- or *phnAc*-type sequences (21, 34) or only PAH-RHDα genes from Gram-negative bacteria (3, 13). In other studies, two-primer systems were used to target PAH-RHDα genes of both Gram-positive and Gram-negative bacteria (4, 37). Only one primer system targeting the Rieske gene fragment was described that amplified a small fragment from PAH-RHDα genes from both Gram-negative and Gram-positive bacteria (24). However, the amplicon size was only 78 bp and the primer might also target genes coding for dioxygenases that attack nonpolar aromatic compounds, such as benzene, toluene, and xylene. Therefore, this work aimed to design an improved primer system that targets PAH-RHDα genes from both Gram-positive and Gram-negative bacteria and provides larger amplicon sizes. The novel primer system was tested *in silico* and validated by sequencing cloned PAH-RHDα genes amplified from total-community (TC) DNA and was used in endpoint and quantitative real-time PCR (qPCR) formats. The primer system was also applied to study the responses of soil microbial communities in two different soils (a Cambisol and a Luvisol representing typical arable soils in Central Europe with different texture compositions) to artificial phenanthrene pollution.

Material and methods

Primers. Primers for PCR amplification of PAH-RHDα genes from both Gram-positive and Gram-negative bacteria were designed based on all 40 PAH-RHDα genes deposited in GenBank (as of 18 December 2007). The criteria for primer design were as follows: (i) no mismatches for the last 7 base pairs at the 3’ end of the primer and (ii) two to six mismatches at the 5’ end. The selected primer system consisting of the 20-mer degenerated forward primer PAHRHDα-396F (5’-ATT GCG CTT AYC AYG GBT GG-3’) and the 21-mer degenerated reverse primer PAH-RHDα-696R (5’-ATA GGT GTC TCC AAC RAA RTT-3’) should allow amplification of an approximately 320-bp fragment from all 40 PAH-RHDα genes. The primer system was tested with *Pseudomonas putida* KT2442 pNF142 and cloned *nahAc*, *phnAc*, *ndo-3*, *ndo-5*, and *nagAc* genes recently described by Gomes et al. (13 and 14) (see Table S1 in the supplemental material). On 26 April 2010, 281 sequences encoding putatively PAH-degradative proteins were obtained using BLAST-N searches whichserved for *in silico* analysis of the primers (see Fig. S1 in the supplemental material).

PCR and qPCR conditions. Optimization of the PCR was done with a gradient of annealing temperatures (46 to 60°C) with DNA from *P. putida* pNF142 and one environmental sample as a target DNA. PCR amplification was performed in a 25 µl reaction mixture consisting of Stoffel buffer (Applied Biosystems, Foster, CA), 0.2 mM deoxynucleoside triphosphates, 2.5 mM MgCl2, 4% (vol/vol) acetamide, 0.1 µg/µL bovine serum albumin, 0.2 µM forward primers, 0.6 µM reverse primers, and 2.5 U Amplitaq DNA polymerase (Stoffel fragment; Applied Biosystems). Denaturation was carried out for 5 min at 94°C with 5 cycles of preamplification, 1 min of denaturation at 94°C, annealing at 46°C for 2 min, and extension at 72°C for 1 min, followed by 30 cycles of 1 min at 95°C, 1 min at 58.5°C, and 1 min at 72°C and 10 min of extension at 72°C. qPCR was performed in 50 µl reaction volumes consisting of Stoffel buffer (Fermentas, St. Leon-Rot, Germany), 0.2 mM deoxynucleoside triphosphates, 2.5 mM MgCl2, 4% (vol/vol) acetamide, 0.1 µg/µL bovine serum albumin, 0.2 µM forward primers, 0.6 µM reverse primers, EvaGreen solution (Biotium, Hayward, CA), and 2.5 U TrueStart Hot Start *Taq* DNA polymerase (Fermentas, St. Leon-Rot, Germany). The amplification was performed as follows. Denaturation was carried out for 5 min at 94°C; 5 cycles of preamplification were done as previously described, followed by 40 cycles of 1 min at 95°C, 45 s at 58.5°C, and 2 min at 60°C and 10 min of extension at 72°C. At the end, dissociation curve analysis was performed in autoincrements from 60°C to 94°C. The standard for quantification of PAH-RHDα genes by qPCR was prepared as follows. The cloned *phnAc* gene fragment (AF061872) was amplified with the primer pair SP6 and T7. The PCR product was gel purified with a Qiaex II Gel Extraction Kit (Qiagen, Hilden, Germany), and the DNA concentration was measured with a BioPhotometer (Eppendorf, Hamburg, Germany). The copy number of the standard DNA was calculated according to the size (920 bp), assuming a molecular mass of 660 Da for a base pair. The standard stock solution of $1.91*10^{10}$ copies was prepared. Primers and a TaqMan probe to quantify 16S rRNA genes by qPCR were previously described (30). PCR amplifications were performed in a 50 µl reaction volume containing 1.25 U TrueStart polymerase (Fermentas, St. Leon-Rot, Germany), 0.2 mM each deoxynucleoside triphosphate, 2.5 mM MgCl2, and 0.25 µM primers and probe. Thermocycles were 10 min at 94°C and 40 cycles consisting of 15 s at 95°C, 15 s at 50°C, and 60 s at 60°C. Templates to generate standard curves were prepared by serial dilutions of gel-purified PCR products from *Escherichia coli* 16S

rRNA genes. The amplification efficiency was tested with four serial dilutions of the Sp6/T7 amplicons obtained from five different cloned RHDα genes (AF061872, GQ479194, GQ479076, GQ479118, and GQ479093) used as templates.

Soil samples. The Eutric Cambisol (10) was collected from a long-term field experiment near Ultuna (60°N, 17°E) in central Sweden. The physicochemical characteristics, as well as previous treatments, have been described (12). The Luvisol (48°N, 11°E) was collected from a farm in Scheyern located north of München, Germany, without a known contamination history. The physicochemical parameters determined in this study for both soils are given in Table 2.1. Briefly, all soils were passed through a 2-mm mesh sieve to remove stones and plant debris and were stored at 4°C for less than 1 month before the batch experiment. The water contents of the soils were determined by leaving the soil in the oven at 110°C overnight. The carbon and nitrogen content of the soils was determined with a Euro EA elemental analyzer, and the pH was measured in 0.01 M $CaCl_2$. Texture was determined after H_2O_2 oxidation of organic matter by sieving and by measuring the X-ray absorption of the soil-water suspension during sedimentation of the soil particles with a Micromeritics Sedigraph 5100 (Micromeritics, Norcross, GA, USA). In addition, phenanthrene sorption was investigated in batch sorption experiments according to the guidelines of the Organization for Economic Cooperation and Development (OECD) (25). The phenanthrene concentrations of the liquid phase were determined by gas chromatography-mass spectrometry (GCMS), and sorption isotherms were calculated using the solubility-normalized Freundlich isotherm (see the supplemental material).

Table 3.1: Properties of soils in this study

Soil parameters		Cambisol	Luvisol
Soil type	Sand (%)	19	16
	Silt (%)	44	70
	Clay (%)	37	14
pH		5.8 ± 0.1	5.5 ± 0.1
Total organic C (mg g^{-1})		14.6 ± 0.2	14.9 ± 0.2
Total N (mg g^{-1})		1.51 ± 0.03	1.59 ± 0.03
C/N ratio		11.3	10.9
depth		0-20 cm	0-20 cm

Experimental design and TC DNA extraction. Prior to the experiment, the soils were kept at room temperature for one week to equilibrate. Contaminating the soils with phenanthrene was done as follows. Twenty grams of each replicate were used as seeding soil, which was first contaminated by spiking it with phenanthrene dissolved in acetone (200 mg ml^{-1}) to reach a final concentration of 20 mg/g soil. Soil that received only acetone was used as seeding soil for the control. The seeding soil was kept in a chemical hood overnight to allow evaporation of the acetone. The adjusted weight of the seed soil was added to 150 g natural soil (final phenanthrene concentration, 2 mg/g soil), mixed thoroughly, and used to fill 250-ml glass flasks that were covered with lids and incubated in the dark at room temperature (23°C). The moisture of the soil was adjusted to ca. 60% of the maximum water-holding capacity by adding sterilized Milli-Q water. At days 0, 21, and 63, samples were taken from the four soil microcosms per treatment and kept at -20°C before DNA extraction. The following abbreviations are used in the text for the treatments: CA, Cambisol control; CP, phenanthrene-spiked Cambisol; LA, Luvisol control; and LP, phenanthrene-spiked Luvisol, taken at day 21 (T21) and day 63 (T63) of the incubation experiment. TC DNA was extracted from 0.5 g of soil after a harsh cell lysis step using the FastPrep FP120 bead-beating system (Qbiogene, Carlsbad, CA, USA) for 30 s at high speed (this step was repeated twice) by means of the Bio-101 DNA spin kit for soil (Qbiogene, Heidelberg, Germany) according to the instructions of the manufacturer.

Cloning, sequencing, and sequence analysis. The PAH-RHDα gene amplicons from four replicates of each soil were combined, gel purified, ligated into pGEM-T vectors, and transformed into competent cells (*E*. *coli* JM109; Promega, Mannheim, Germany) according to the instructions of the manufacturer. After the primer sequences were removed, the sequences were analyzed by BLAST-N and TBLASTX (http://www.ncbi.nlm.nih.gov/BLAST/) to identify PAH-RHDα gene sequences (ca. 280 bp). All sequences obtained were submitted to GenBank (see Table S2 in the supplemental material). Translated amino acid sequences of cloned PAH-RHDα genes, as well as of most known PAHRHDα groups, were aligned, and the tree was calculated according to the neighbor joining method and tested by bootstrap analysis using Molecular Evolutionary Genetics Analysis (MEGA4) (20) integrated software.

Based on a DNA distance matrix created with Phylip, the program DOTUR (distance-based operational taxonomic unit [OTU] and richness determination) (27) was used to assign sequences to operational units (Ous) (sequences that have less than 5% DNA distance) and for rarefaction analyses. Rarefaction curves were plotted and further analyzed with three regression equations: (regression 1 [R1]) $y = a*(1-\exp^{-bx})$, (R2) $y=a*(1- \exp^{-bx\hat{}c)}$ (31), and (R3) $y= (a*x)/ (b*x)$, where x is the sample size, y is the observed number of Ous, a is the number of Ous expected with infinite sample size, and b and c are constants to fit the model. SigmaPlot v. 11.0 was used for regression analysis.

Phenanthrene analysis. Quantitation of phenanthrene was done according to the method of Baran and Oleszczuk (2) with the following modifications. Briefly, 1 gram soil was extracted using an ultrasonic bath and overhead shaker. The final extract was separated with an acetonitrile-water gradient on a C18 column (Luna C18 [2]; 100 A, 150 by 2.0 mm, 3 mm; Phenomenex, Aschaffenburg, Germany) and the phenanthrene was detected at 254 nm (UVD 340 S UV detector; Dionex). External quantitation was done with the multicomponent standard solution SRM 1647D (National Institute of Standards and Technology, Promochem, Wesel, Germany) at concentrations of 1.0, 2.5, 5.0, and 7.5 µg/ml. Linearity was excellent (R^2 = 0.996).

RESULTS

Design of a new PAH-RHDα primer set. *In silico* analysis indicated that the primer set designed on the basis of all 40 different PAH-RHDα sequences deposited in GenBank (as of December 2007) would amplify a wide spectrum of PAH-RHDα genes, such as *nahAc*, *phnAc*, *nagAc*, *bphA1*, *ndo-3*, and *ndo-5* from Gram-negative bacteria and *phdAc*, *narAa*, *nidA*, and other RHDα genes from Gram-positive bacteria, as no mismatches were detected for at least 7 base pairs from the 3' ends of both the forward and reverse primers (see Fig. S2 in the supplemental material). Recently, 281 sequences putatively coding for PAH-degradative genes from GenBank (as of 26 April 2010) were obtained using BLAST-N searches and used for *in silico* analysis of the primers. Of the 281 sequences, only 143 were long enough to be checked for the specificity of both primers; 138 sequences had 7-bp perfect matches at the 3- ends of both primers, while 2 to 6 mismatches were observed at

the 5- ends of the primers. However, 130 out of 281 sequences in GenBank were too short for *in silico* analysis of both primers, and thus, no prediction for amplification was possible. However, the annealing site that could be analyzed met the criteria described above (see Fig. S1 in the supplemental material). In order to amplify diverse PAH-RHDα sequences from TC DNA despite mismatches of the primers at the 5' ends, a preamplification with five cycles using an annealing temperature of only 46°C for 2 min was done, followed by 30 cycles with an annealing temperature of 58.5°C. Furthermore, amplicons of the expected size (approximately 320 bp) were obtained from the genomic DNA of *P. putida* pNF142 and of eight cloned *ndo* genes (see Table S1 in the supplemental material). The same primer set was also used for qPCR. Testing the amplification efficiencies of five different RHDα genes revealed that the amplification efficiencies ranged from 73.3% to 97.4% (see Table S3 in the supplemental material). The cloned *phnAc* gene used for the standard curve had the lowest amplification efficiency.

Detection of PAH-RHDα genes in soils. The new primer set was used to study the occurrence and diversity of PAH-RHDα genes in two different soils (Luvisol and Cambisol) and variants that were spiked with 2 g phenanthrene per kg soil. At day 21, amplicons were detected only in all four replicates of the phenanthrene contaminated Luvisol. At day 63, amplicons of putative PAH-RHDα genes were amplified from TC DNA of all replicates of both contaminated soils, but not from both control soils (data not shown).

Diversity of PAH-RHDα genes in two phenanthrene-contaminated soils. To experimentally determine the specificity of the novel primers and the diversity of PAH-RHDα genes amplified with this system from soil DNA, cloning of the pooled PCR products and sequencing were done. BLAST-N and TBLASTX (http://www.ncbi.nlm.nih.gov/BLAST/) were used to identify the PAH-RHDα gene sequences (ca. 280 bp). Thirty-three and 37 sequences analyzed from the clone library from phenanthrene-contaminated Luvisol at day 21 (T21LP) and day 63 (T63LP), respectively, were identified as PAH-RHDα gene sequences corresponding to 61.1% and 94.9% of the clones with sufficient read length (>233 bp). All 33 sequences from clone library T21LP and 34 out of 37 sequences from clone library T63LP shared 98% of the *phnAc* gene sequence of *Burkholderia* sp. Strain Eh1-1

(Fig. 1; see also Table S4 in the supplemental material). Two sequences were affiliated with the PAH-RHDα genes of *Mycobacterium* sp. Strain JLS. These genes might be novel, as they shared only 87% nucleotide sequence similarity. One of the 37 sequences of the T63LP library had 100% similarity to one PAH-RHDα gene of the *Novosphingobium aromaticivorans* DSM 12444 plasmid pNL1.

The diversity of PAH-RHDα genes in the phenanthrene-spiked Cambisol could not be determined for day 21, as no amplicon was obtained. The vast majority (85/94) of T63CP sequences with sufficient read length were identified as PAH-RHDα genes. Many of the PAH-RHDα gene sequences determined (62/85) are likely to be novel, as they shared only 72 to 90% similarity (BLAST-N) with their closest corresponding genes in GenBank. However, 18 out of 87 sequences were affiliated with different PAH-RHDα genes from mycobacteria with more than 97% nucleotide sequence identity. Two sequences of clones showed high similarity to PAH-RHDα genes of *Burkholderia* spp. The phylogenetic analysis of the PAH-RHDα genes amplified from T63CP revealed high diversity, with a Chao1 value of 30.2 (Table 2), and confirmed that the novel primer set amplified PAH-RHDα genes from both Gram-positive and Gram-negative bacteria. The majority of PAH-RHDα genes amplified were affiliated with PAH-RHDα genes from *Mycobacterium* spp. (31/85) and two PAH-RHDα genes from *N. aromaticivorans* DSM 12444 plasmid pNL1 (50/ 85) (Fig. 1; see also Table S4 in the supplemental material).

Observed Ous and regression analysis revealed that a higher diversity of PAH-RHDα genes was enriched in the contaminated Cambisol than in the Luvisol. A total of 23 Ous were identified in the clone library of the Cambisol (T63CP; 85 sequences analyzed) but only four Ous in the clone library of the Luvisol (T63LP; 37 sequences analyzed). To compare the richnesses of the two clone libraries, 37 sequences were sampled randomly 1,000 times from the clone library of T63CP, and an average of 16.4 Ous was reached (see Fig. S3 in the supplemental material). Regression analyses done with three equations suggested a significantly higher number of Ous from T63CP than from T63LP was expected with infinite sample size. According to regression analysis, the two clone libraries generated from amplicons obtained from Luvisol and Cambisol at day 63 after phenanthrene pollution covered 80 to 100% and 74 to 100% of PAH-RHDα gene diversity, respectively. Clone library T63LP had a

higher coverage of diversity than T63CP, as suggested by regression 2 and regression 3 (Table 2).

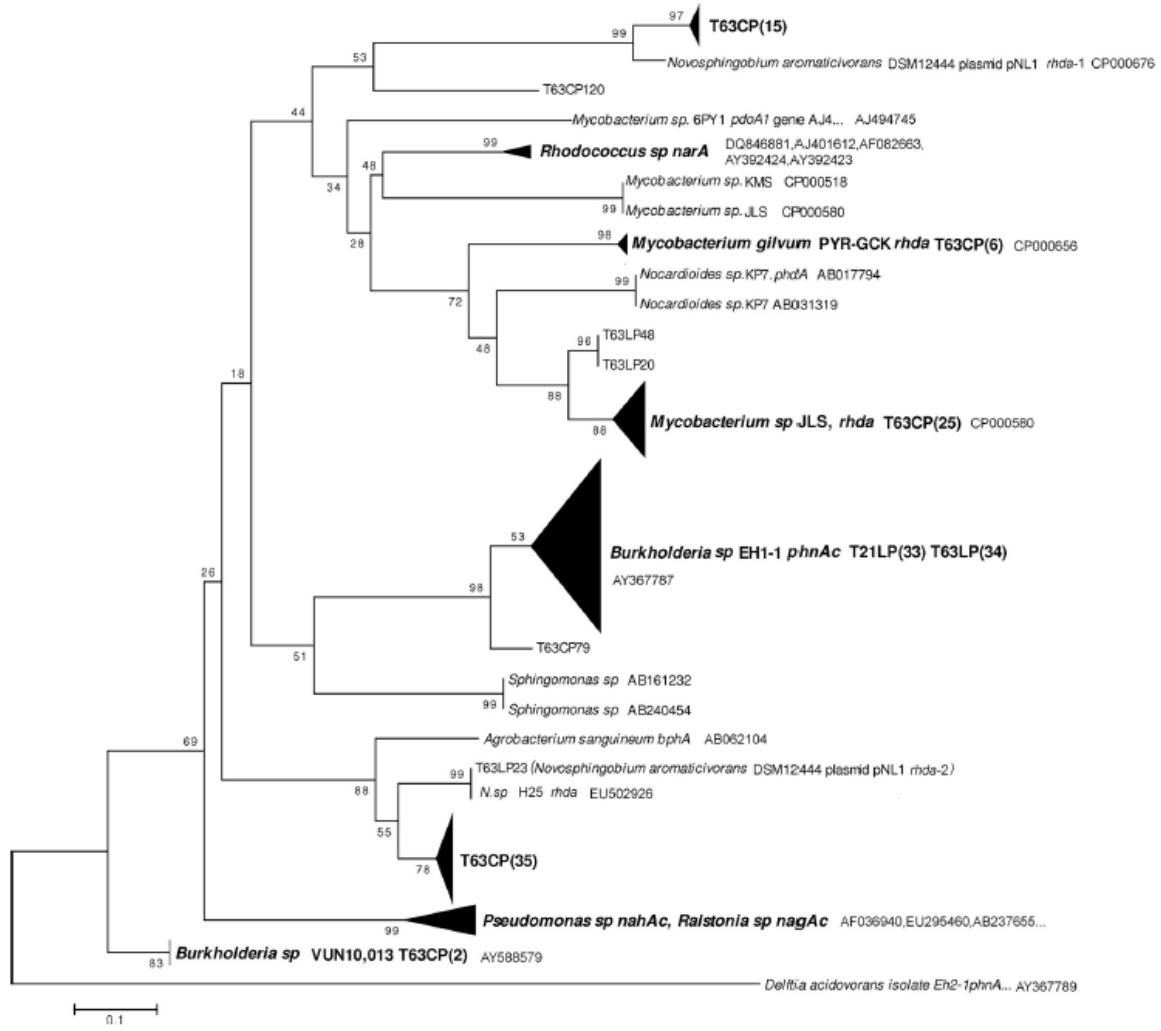

FIG. 3.1. Neighbor-joining rooted phylogenetic tree based on the multiple alignments of PAH-RHDα gene sequences from translated amino acid sequences. The triangles represent compressed branches containing various sequences from different treatments as indicated; the vertical length of a triangle reflects the number of sequences, and the horizontal length reflects the largest distance between sequences. Value at each node = (bootstrap value/500) * 100.

Quantification of PAH-RHDα genes. Total bacterial 16S rRNA gene copy numbers and copies of PAH-RHDα genes were determined by qPCR for all samples collected at day 63. In the Luvisol, a significantly higher 16S rRNA gene copy number per gram soil was detected in phenanthrene-contaminated soil than in the control soil. Also, the phenanthrene-contaminated Cambisol showed a slightly higher 16S rRNA gene copy number than the control. Between the two control soils, there was no significant difference in 16S rRNA gene copy numbers, while the contaminated Luvisol had significantly higher copy numbers than the contaminated Cambisol. PAH-

RHDα genes were detected only in all replicates of both contaminated soil samples, with 2.03*107±1.05*107 copies per gram soil and $1.68*10^6 \pm 0.66*10^6$ copies per gram soil for T63LP and T63CP, respectively, but not in the corresponding control soils (Fig. 2).

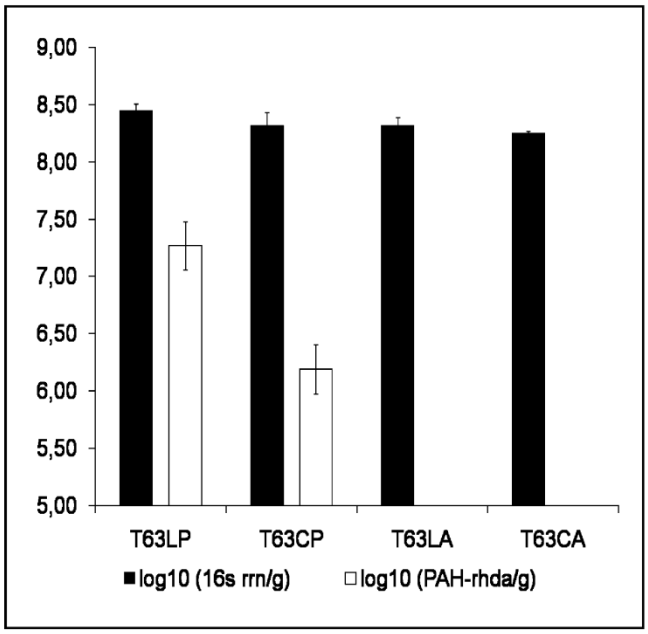

FIG. 3.2. Determination of 16S rRNA gene and PAH-RHDα gene copy numbers for samples collected after 63 days of incubation by qPCR. The error bars indicate standard deviations of the four replicates of each treatment.

Detection of phenanthrene in the Luvisol and in the Cambisol. Concentrations of phenanthrene in both soils were followed at days 0, 7, 21, and 63 after spiking. No detectable phenanthrene was present in the Luvisol in the control set; small amounts (1.7 and 1.8 mg/kg) were detected from two samples of the Cambisol controls. Seven days after spiking, the amount of phenanthrene extracted from the Cambisol, but not from the Luvisol, decreased. In the soil samples taken after 21 days, the amounts of phenanthrene extracted from both soils were similar, but only about 42% of the initial concentration in the soil was detectable. At day 63, no phenanthrene was extractable from the Luvisol while 49±52 mg/kg was still detected in the Cambisol (Fig. 3).

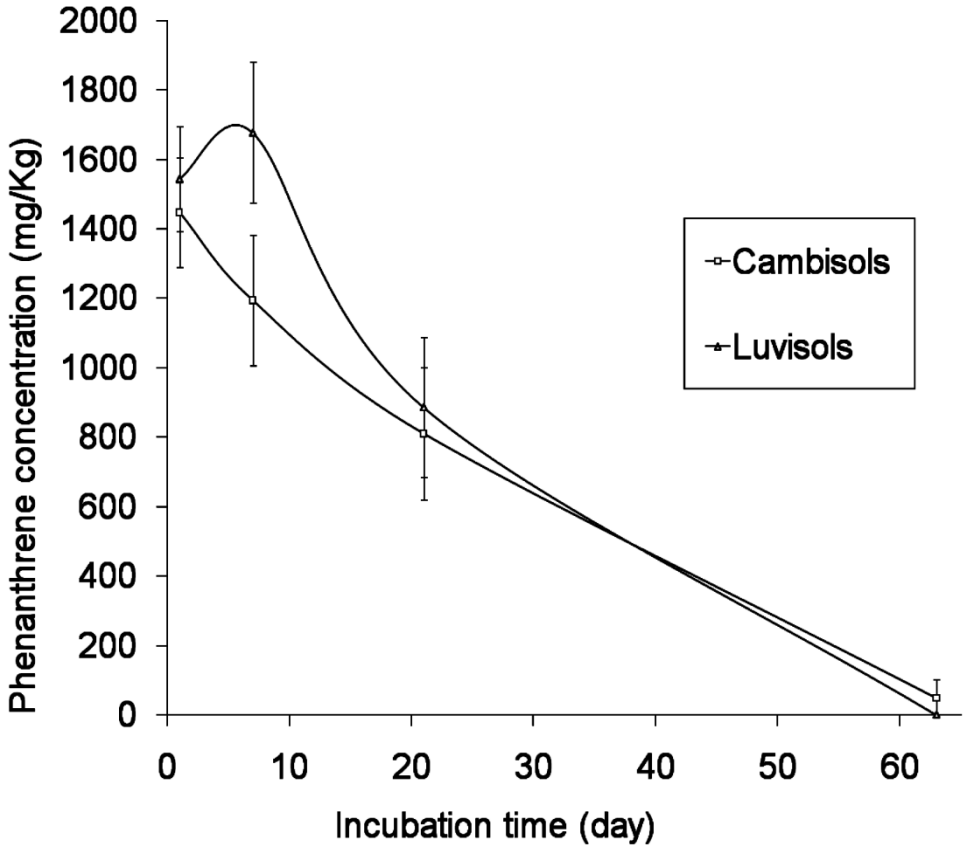

FIG. 3.3. Phenanthrene concentrations in the soil at different time points during incubation. The error bars indicate standard deviationsof the four replicates of each treatment.

DISCUSSION

Genes coding for ring-hydroxylating dioxygenases are highly diverse, and thus, two important points were considered in order to develop primers to successfully amplify a wide range of RHDα genes from mixed template DNA. No mismatches in the last 7 bp at the 3' end of each primer but 2 to 6 mismatches at the 5' end were allowed. More importantly, to achieve amplification from mixed template DNA despite the mismatches at the 5' end, a preamplification step with five cycles at low annealing temperature (46°C) was essential to generate templates that perfectly matched the primers. In contrast to other studies, primer validation was mainly based on sequencing cloned PAH-RHDα gene amplicons from two phenanthrene-contaminated soils. Sequence analysis confirmed not only the amplification of PAH-RHDα genes from Gram-positive and Gram-negative bacteria, but also high specificity of the primer system, as more than 90% of the sequences obtained for

T63CP and T63LP were affiliated with PAH-RHDα genes. Only sequences obtained from the clone library of T21LP showed a rather high proportion of sequences (39%) that had no similarity to PAH-RHDα genes in GenBank. The amount of amplicon DNA obtained from T21LP was just visible on the agarose gel, indicating a low number of target sequences. Also, other authors reported increased proportions of nonspecific amplicons when PCR amplification was done from DNA with a low concentration of target genes (3). Without exposure to phenanthrene, no PAH-RHDα gene amplicons were detected by PCR and qPCR, and thus, the copy number was below the estimated detection limit of 10^4 copies per g of soil. Other studies using different primer systems also reported the absence of PCR amplicons when TC DNA from non-contaminated soils or groundwater was used (3, 9, 14, 35). In contrast to the only previously published primer system (24) that also targeted RHDα genes of Gram-positive and Gram-negative bacteria, the amplified sequence that can be used for sequence analysis is much longer (approximately 280 bp compared to approximately 47 bp) and the primers specifically target PAH-RHDα genes.

Upon spiking Luvisol or Cambisol with phenanthrene, the abundance of PAH-RHDα gene-carrying populations must have increased, as PCR products were obtained only from these soil samples. Furthermore, the increased abundance of soil bacteria carrying RHDα genes indicated their involvement in the degradation of phenanthrene. Also, in other studies, the relative abundance of PAH-RHDα genes was correlated with the level of PAH contamination (4, 9, 14, 35). Microbial communities inhabiting the two soils studied showed strikingly different abundances and diversities of PAH-RHDα genes after exposure to the same concentrations of phenanthrene under identical incubation conditions. The diversity of PAH-RHDα genes in Luvisol was rather low, and all sequences from the T21LP clone library of PAH-RHDα gene amplicons had the highest nucleotide identity with the *phnAc* gene of the *Burkholderia* strain Eh-1, a strain that was originally isolated from a phenanthrene enrichment of a contaminated soil from a former coal gasification plant in Wisconsin (33). These amplicons were obtained even though sequences belonging to this *phnAc* gene cluster have 2 to 4 mismatches with the forward primer (PAH-RHDα-396f) and 4 to 5 with the reverse primer (PAH-RHDα-696r) (see Fig. S1 in the supplemental material). While the clone library of T63LP was still dominated by *phnAc*, a completely different situation was encountered for the PAH-RHDα gene library obtained for the Cambisol. This clone library had much higher diversity (Fig. 1

and Table 2; see Table S4 in the supplemental material), and the PAH-RHDα genes detected were derived from high-GC Gram-positive bacteria (*Mycobacterium* and *Rhodococcus*), *Alphaproteobacteria* (*Agrobacterium* and *Novosphingobium*), and *Betaproteobacteria* (*Burkholderia* and *Acidovorax*). Although several of the sequences showed high sequence similarities to known PAH-RHDα gene sequences, many sequences had only rather low similarities (BlastN and BlastX) to sequences in GenBank, indicating novel genes not yet described. Exposure to phenanthrene in the Luvisol resulted in more than one order of magnitude higher PAH-RHDα gene copy numbers in the Luvisol than in the Cambisol detected at day 63. Whether the initial abundance of PAH-RHDα genes in the Luvisol was already higher at the beginning of the experiment or whether the differences were due to the more rapidly growing *phnAc*-carrying population remains unclear, as in both soils, the abundance of PAH-RHDα genes was below the detection limit. Adaptation of soil microbial communities to PAHs is assumed to occur through induction of enzymes involved in the biodegradation of PAHs or an increase in the number of bacteria with degradative capacities (23). The high abundance of a fast growing population carrying the *phnAc* gene in the Luvisol might have prevented sufficient growth of slower-growing populations carrying PAH-RHDα genes and their detection in the Luvisol library. In contrast to the Luvisol, the response of the bacterial communities in the Cambisol to phenanthrene was delayed, as amplicons were obtained only at day 63 but a much higher diversity of PAH-RHDα genes was observed. A high proportion of the genes (36.4%) that were detected in response to phenanthrene were affiliated with PAH-RHDα genes from mycobacteria. Nine of the sequences displayed high similarity to genes identified on the recently completed genome sequence of *Mycobacterium gilvum* PYR-GCK (7, 19). Recently, Fredslund et al. (11) demonstrated surface sliding motility of *M. gilvum* and discussed the implications for PAH degradation in contaminated soil. Furthermore, PAH-degrading mycobacteria were shown to be enriched on PAH-containing soil components (32).

The differences in the responses of the bacterial communities of both soils to phenanthrene contamination might also be related to the different soil properties (Table 1). The composition and heterogeneity of soil organic matter can greatly influence the mineralization of PAHs (5, 23, 26). However, the phenanthrene sorption experiments performed in this study (see Fig. S4 in the supplemental material) showed no significant difference in phenanthrene sorption between the two soils.

This implies that the difference in clay content had no effect on phenanthrene sorption and that the organic matter content was probably the only factor determining phenanthrene sorption. The two soils have similar C/N ratios and land use histories and therefore probably similar organic matter compositions. On the other hand, the differences in soil properties are assumed to be an important factor shaping soil microbial communities. Soil type-dependent differences in the bacterial community composition are well documented (28).

In conclusion, the bacterial community compositions present in the two soils, as well as soil texture-related factors (e.g., porosity or the matrix potential), might have contributed to the different responses. Furthermore, the batch experiment also showed that many degradative genes will remain undiscovered in soil metagenomics unless enrichments are performed to increase the abundance of rare populations carrying these genes. Thus, it is important to keep in mind that PCR based detection with a novel primer system will also provide insights into the degradative potentials of bacterial populations that occur with 104 or more gene copies per gram of soil. Thus, soil enrichments could be an important means to uncover the biodegradative potentials of uncommon populations.

ACKNOWLEDGMENTS This work was supported by DFG SPP1315 (SM59/8–1, SP255/19–1, and KO1035/33–1).

We thank Ilse-Marie Jungkurth for carefully checking our manuscript.

REFERENCES

1. **Amann, R. I., W. Ludwig, and K.-H. Schleifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59:**143–169.

2. **Baran, S., and P. Oleszczuk.** 2002. Chromatographic determination of polycyclic aromatic hydrocarbons (PAH) in sewage sludge, soil, and sewage sludge-amended soils. Pol. J. Environ. Studies **11:**609–615.

3. **Bordenave, S., M. Goni-Urriza, C. Vilette, S. Blanchard, P. Caumette, and R. Duran.** 2008. Diversity of ring-hydroxylating dioxygenases in pristine and oil contaminated microbial mats at genomic and transcriptomic levels. Environ. Microbiol. **10:**3201–3211.

4. **Ce´bron, A., M. P. Norini, T. Beguiristain, and C. Leyval.** 2008. Real-time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHD_) genes from Gram positive and Gram negative bacteria in soil and sediment samples. J. Microbiol. Methods **73:**148–159.

5. **Celis, R., H. De Jonge, L. W. De Jonge, M. Real, M. C. Hermosin, and J. Cornejo.** 2006. The role of mineral and organic components in phenanthrene and dibenzofuran sorption by soil. Eur. J. Soil Sci. **57:**308–319.

6. **Chain, P. S., V. J. Denef, K. T. Konstantinidis, L. M. Vergez, L. Agullo, V. L. Reyes, L. Hauser, M. Cordova, L. Gomez, M. Gonzalez, M. Land, V. Lao, F. Larimer, J. J. LiPuma, E. Mahenthiralingam, S. A. Malfatti, C. J. Marx, J. J. Parnell, A. Ramette, P. Richardson, M. Seeger, D. Smith, T. Spilker, W. J. Sul, T. V. Tsoi, L. E. Ulrich, I. B. Zhulin, and J. M. Tiedje.** 2006. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. Proc. Natl. Acad. Sci. U. S. A. **103:**15280–15287.

7. **Dean-Ross, D., and C. E. Cerniglia.** 1996. Degradation of pyrene by *Mycobacterium flavescens*. Appl. Microbiol. Biotechnol. **46:**307–312.

8. **Debruyn, J. M., and G. S. Sayler.** 2009. Microbial community structure and biodegradation activity of particle-associated bacteria in a coal tar contaminated creek. Environ. Sci. Technol. **43:**3047–3053.

9. **Flocco, C. G., N. C. Gomes, C. W. Mac, and K. Smalla.** 2009. Occurrence and diversity of naphthalene dioxygenase genes in soil microbial communities from the maritime Antarctic. Environ. Microbiol. **11:**700–714.

10. **Food and Agriculture Organization of the United Nations.** 2006. World soil resources reports. FAO, Rome, Italy.

11. **Fredslund, L., K. Sniegowski, L. Y. Wick, C. S. Jacobsen, R. De Mot, and D. Springael.** 2008. Surface motility of polycyclic aromatic hydrocarbon (PAH)-degrading mycobacteria. Res. Microbiol. **159:**255–262.

12. **Gerzabek, M. H., R. S. Antil, I. Kogel-Knabner, H. Knicker, H. Kirchmann, and G. Haberhauer.** 2006. How are soil use and management reflected by soil organic matter characteristics: a spectroscopic approach. Eur. J. Soil Sci. **57:**485–494. 4770 DING ET AL. APPL. ENVIRON. MICROBIOL.

13. **Gomes, N. C., L. R. Borges, R. Paranhos, F. N. Pinto, E. Krögerrecklenfort, L. C. Mendonca-Hagler, and K. Smalla.** 2007. Diversity of *ndo* genes in mangrove sediments exposed to different sources of polycyclic aromatic hydrocarbon pollution. Appl. Environ. Microbiol. **73:**7392–7399.

14. **Gomes, N. C., I. A. Kosheleva, W. R. Abraham, and K. Smalla.** 2005. Effects of the 75luorine75s strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. FEMS Microbiol. Ecol. **54:**21–33.

15. **Habe, H., and T. Omori.** 2003. Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. Biosci. Biotechnol. Biochem. **67:**225–243.

16. **Johnsen, A. R., L. Y. Wick, and H. Harms.** 2005. Principles of microbial PAH-degradation in soil. Environ. Pollut. **133:**71–84.

17. **Kanaly, R. A., and S. Harayama.** 2000. Biodegradation of high-molecular weight polycyclic aromatic hydrocarbons by bacteria. J. Bacteriol. **182:**2059–2067.

18. **Kim, S. J., O. Kweon, R. C. Jones, R. D. Edmondson, and C. E. Cerniglia.** 2008. Genomic analysis of polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. Biodegradation **19:**859–881.

19. **Kim, Y. H., K. H. Engesser, and C. E. Cerniglia.** 2005. Numerical and genetic analysis of polycyclic aromatic hydrocarbon-degrading mycobacteria. Microb. Ecol. **50:**110–119.

20. **Kumar, S., M. Nei, J. Dudley, and K. Tamura.** 2008. MEGA: a biologistcentric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. **9:**299–306.

21. **Lloyd-Jones, G., A. D. Laurie, D. W. F. Hunter, and R. Fraser.** 1999. Analysis of catabolic genes for naphthalene and phenanthrene degradation in contaminated New Zealand soils. FEMS Microbiol. Ecol. **29:**69–79.

22. **Lozada, M., J. P. Riva Mercadal, L. D. Guerrero, W. D. Di Marzio, M. A. Ferrero, and H. M. Dionisi.** 2008. Novel aromatic ring-hydroxylating dioxygenase genes from coastal marine sediments of Patagonia. BMC Microbiol. **8:**50.

23. **Macleod, C. J. A., and K. T. Semple.** 2002. The adaptation of two similar soils to pyrene catabolism. Environ. Pollut. **119:**357–364.

24. **Ni Chadhain, S. M., R. S. Norman, K. V. Pesce, J. J. Kukor, and G. J. Zylstra.** 2006. Microbial dioxygenase gene population shifts during polycyclic aromatic hydrocarbon biodegradation. Appl. Environ. Microbiol. **72:**4078–4087.

25. **Organization for Economic Cooperation and Development.** 2000. Test no. 106: adsorption-desorption using a batch equilibrium method. OECD guidelines for testing chemicals. OECD Publications, Paris, France.

26. **Reid, B. J., C. J. MacLeod, P. H. Lee, A. W. Morriss, J. D. Stokes, and K. T. Semple.** 2001. A simple 14C-respirometric method for assessing microbial catabolic potential and contaminant bioavailability. FEMS Microbiol. Lett. **196:**141–146.

27. **Schloss, P. D., and J. Handelsman.** 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. **71:**1501–1506.

28. **Smalla, K., M. Oros-Sichler, A. Milling, H. Heuer, S. Baumgarte, R. Becker, G. Neuber, S. Kropf, A. Ulrich, and C. C. Tebbe.** 2007. Bacterial diversity of soils assessed by DGGE, T-RFLP and SSCP fingerprints of PCR-amplified 16S rRNA gene fragments: do the different methods provide similar results? J. Microbiol. Methods **69:**470–479.

29. **Staley, J. T., and A. Konopka.** 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu. Rev. Microbiol. **39:**321–346.

30. **Takai, K., and K. Horikoshi.** 2000. Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. Appl. Environ. Microbiol. **66:**5066–5072.

31. **Thiel, V., S. C. Neulinger, T. Staufenberger, R. Schmaljohann, and J. F. Imhoff.** 2007. Spatial distribution of sponge-associated bacteria in the Mediterranean sponge *Tethya aurantium*. FEMS Microbiol. Ecol. **59:**47–63.

32. **Uyttebroek, M., P. Breugelmans, M. Janssen, P. Wattiau, B. Joffe, U. Karlson, J. J. Ortega-Calvo, L. Bastiaens, A. Ryngaert, M. Hausner, and D. Springael.** 2006. Distribution of the Mycobacterium community and polycyclic aromatic hydrocarbons (PAHs) among different size fractions of a long-term PAH-contaminated soil. Environ. Microbiol. **8:**836–847.

33. **Vacca, D. J., W. F. Bleam, and W. J. Hickey.** 2005. Isolation of soil bacteria adapted to degrade humic acid-sorbed phenanthrene. Appl. Environ. Microbiol. **71:**3797–3805.

34. **Wilson, M. S., C. Bakermans, and E. L. Madsen.** 1999. In situ, real-time catabolic gene expression: extraction and characterization of naphthalene dioxygenase mRNA transcripts from groundwater. Appl. Environ. Microbiol. **65:**80–87.

35. **Yagi, J. M., and E. L. Madsen.** 2009. Diversity, abundance, and consistency of microbial oxygenase expression and biodegradation in a shallow contaminated aquifer. Appl. Environ. Microbiol. **75:**6478–6487.

36. **Yagi, J. M., D. Sims, T. Brettin, D. Bruce, and E. L. Madsen.** 2009. The genome of *Polaromonas naphthalenivorans* strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. Environ. Microbiol. **11:**2253–2270.

37. **Zhou, H. W., C. L. Guo, Y. S. Wong, and N. F. Tam.** 2006. Genetic diversity of dioxygenase genes in polycyclic aromatic hydrocarbon-degrading bacteria isolated from mangrove sediments. FEMS Microbiol. Lett. **262:**148–157.

Supplements

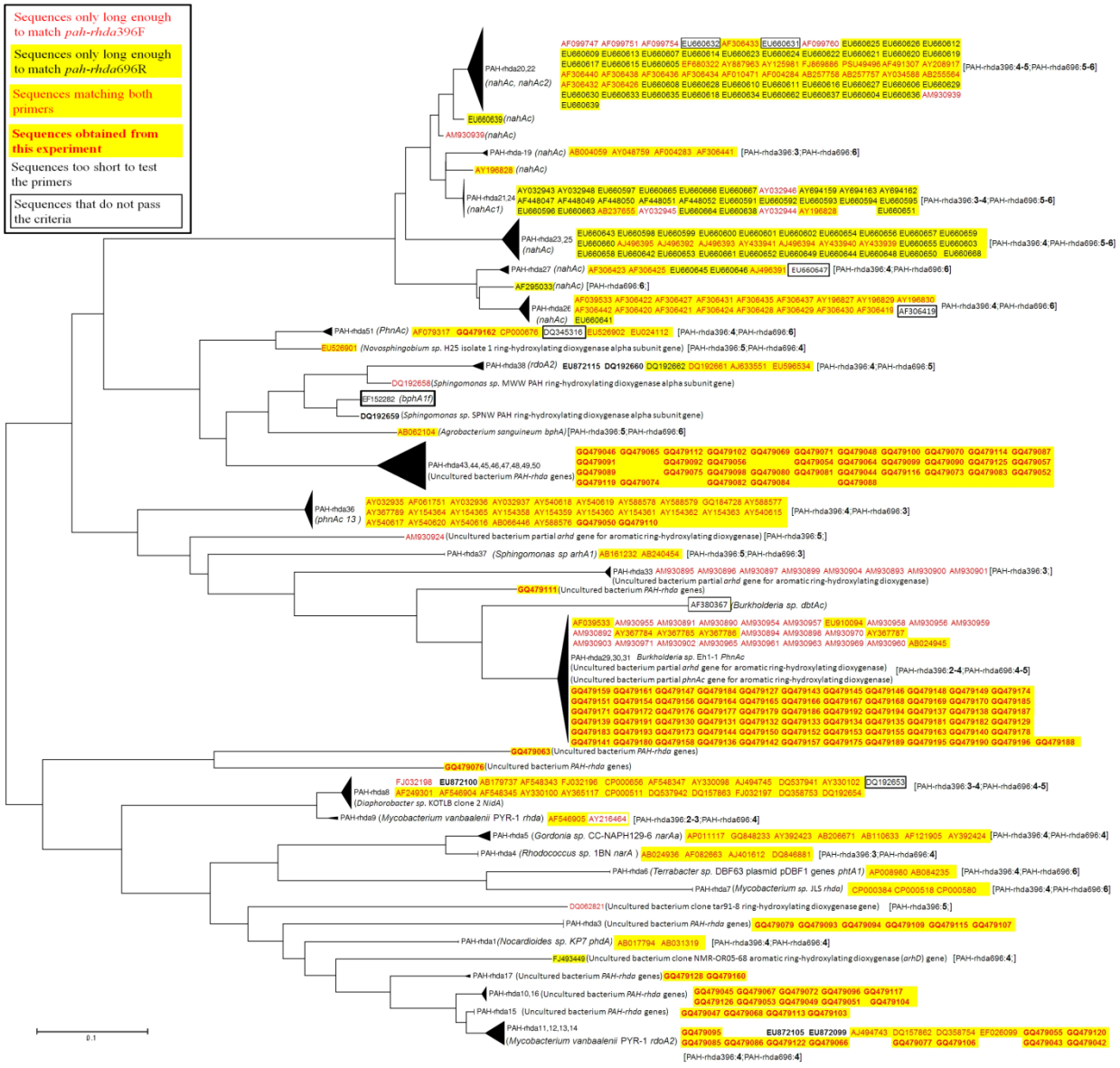

Fig. S1: *In silico* analyses of the primers based on 281 sequences of *pah-rhdα* genes

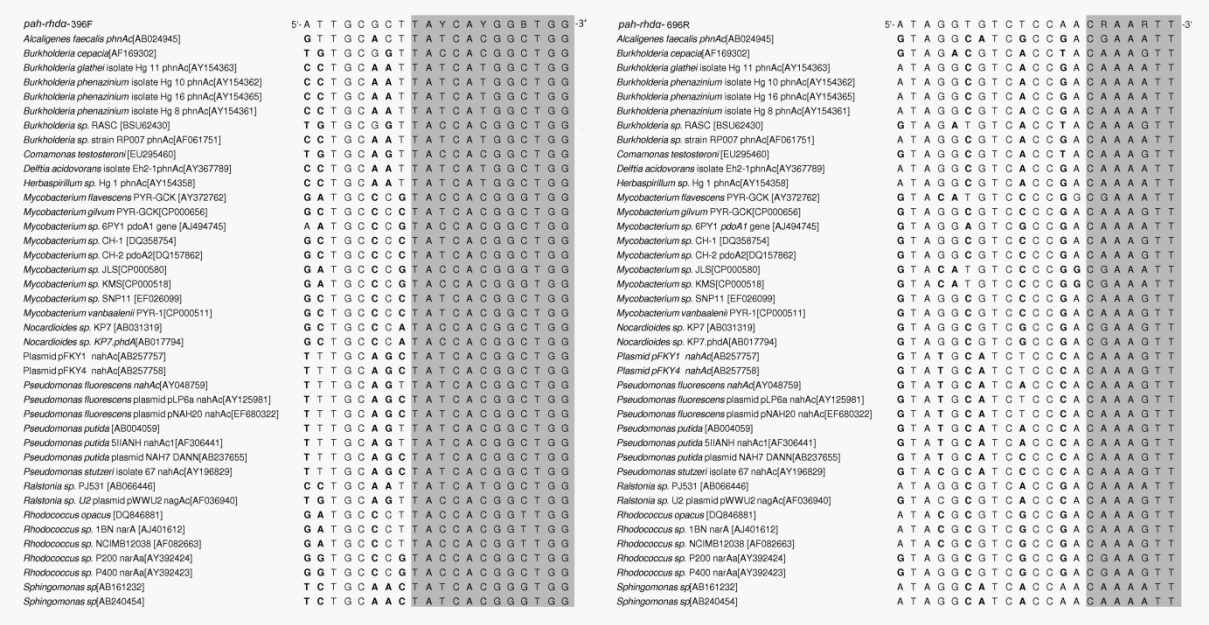

Fig. S2: Alignment of primers and 40 targeted PAH-RHDα genes

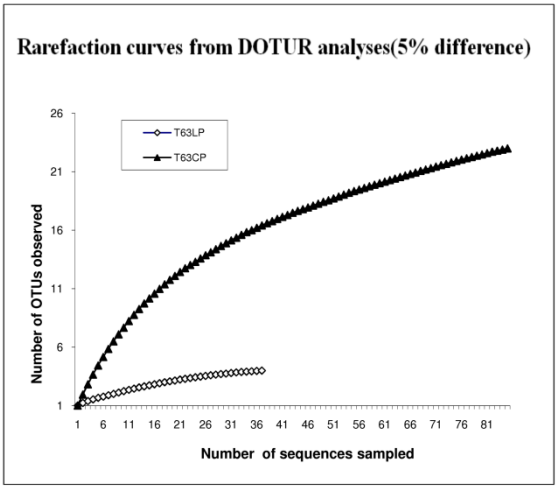

Fig. S3: Plot of the rarefaction curves for clone library T63LP and T63CP

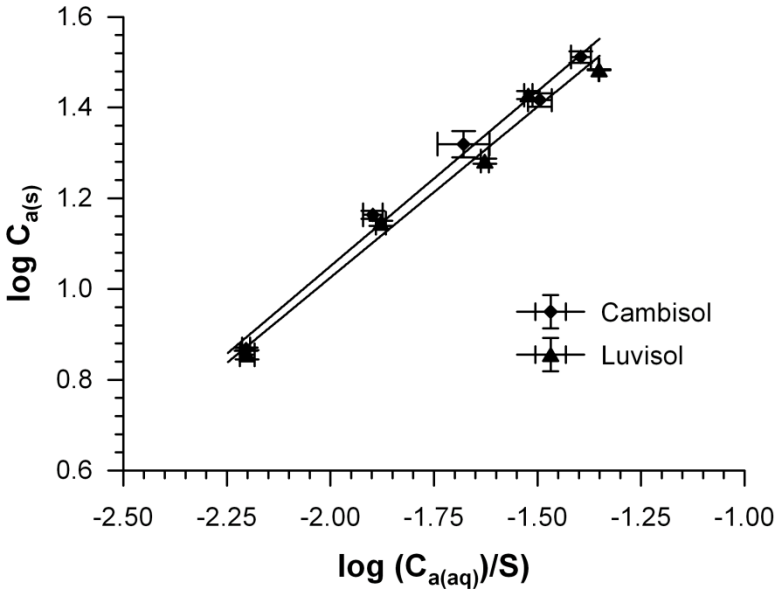

Fig S4: Phenanthrene sorption isotherm. Ca(aq) and Ca(s) are the concentrations of phenanthrene in the solution (aq) and in the sample (s) at equilibrium, S is the super cooled liquid solubility of phenanthrene. Error bars indicate standard deviation of 3 replicates.

Table S1: Strain and cloned *ndo* genes used to test the primers

Gene type	Origin	References
nahAc	*P. putida* KT2442(pNF142)	**(4)**
phnAc AY540615	Cloned fragment (740 bp) of *phn*Ac-gene	**(4)**
*ndo*3 EF455675	Cloned *ndo* fragment (ca 740 bp) from mangroves sediments in Guanabara Bay	**(3)**
*ndo*5 EF455676	Cloned *ndo* fragment (ca 740 bp) from mangroves sediments in Guanabara Bay	**(3)**
nagAc	Cloned *ndo* fragment (ca 740 bp) from mangroves sediments in Guanabara Bay	**(3)**
nahAc EU660594	Cloned *ndo* fragment (ca 740 bp) from PAH polluted Antarctic soil	**(2)**
nahAc EU660600	Cloned *ndo* fragment (ca 740 bp) from PAH polluted Antarctic soil	**(2)**
nahAc EU660650	Cloned *ndo* fragment (ca 740 bp) from PAH polluted Antarctic soil	**(2)**
nahAc EU660620	Cloned *ndo* fragment (ca 740 bp) from PAH polluted Antarctic soil	**(2)**

Table S2: List of the clones and their corresponding accession number.

Clone	Accession	Clone	Accession	Clone	Accession	Clone	Accession
T63CP1	GQ479042	T63CP46	GQ479092	T63LP17	GQ479142	T21LP2	GQ479192
T63CP10	GQ479043	T63CP49	GQ479093	T63LP19	GQ479143	T21LP104	GQ479193
T63CP11	GQ479044	T63CP50	GQ479094	T63LP33	GQ479144	T21LP7	GQ479194
T63CP12	GQ479045	T63CP52	GQ479095	T63LP4	GQ479145	T21LP68	GQ479195
T63CP13	GQ479046	T63CP56	GQ479096	T63LP2	GQ479146	T21LP63	GQ479196
T63CP14	GQ479047	T63CP57	GQ479097	T63LP22	GQ479147	T63CP144	GQ479087
T63CP15	GQ479048	T63CP59	GQ479098	T63LP24	GQ479148	T63CP41	GQ479088
T63CP16	GQ479049	T63CP63	GQ479099	T63LP26	GQ479149	T63CP42	GQ479089
T63CP23	GQ479050	T63CP64	GQ479100	T63LP27	GQ479150	T63CP44	GQ479090
T63CP24	GQ479051	T63CP65	GQ479101	T63LP28	GQ479151	T63CP45	GQ479091
T63CP27	GQ479052	T63CP67	GQ479102	T63LP29	GQ479152	T63LP25	GQ479137
T63CP30	GQ479053	T63CP68	GQ479103	T63LP30	GQ479153	T63LP21	GQ479138
T63CP34	GQ479054	T63CP70	GQ479104	T63LP31	GQ479154	T63LP12	GQ479139
T63CP35	GQ479055	T63CP71	GQ479105	T63LP35	GQ479155	T63LP15	GQ479140
T63CP4	GQ479056	T63CP72	GQ479106	T63LP5	GQ479156	T63LP16	GQ479141
T63CP40	GQ479057	T63CP73	GQ479107	T63LP34	GQ479157	T21LP107	GQ479187
T63CP9	GQ479058	T63CP75	GQ479108	T63LP13	GQ479158	T21LP4	GQ479188
T63CP18	GQ479059	T63CP76	GQ479109	T63LP7	GQ479159	T21LP16	GQ479189
T63CP131	GQ479060	T63CP77	GQ479110	T63LP20	GQ479160	T21LP3	GQ479190
T63CP128	GQ479061	T63CP79	GQ479111	T63LP9	GQ479161	T21LP163	GQ479191
T63CP141	GQ479062	T63CP82	GQ479112	T63LP23	GQ479162	T63CP130	GQ479080
T63CP101	GQ479063	T63CP83	GQ479113	T63LP14	GQ479163	T63CP137	GQ479081
T63CP104	GQ479064	T63CP84	GQ479114	T21LP13	GQ479164	T63CP138	GQ479082
T63CP66	GQ479065	T63CP85	GQ479115	T21LP113	GQ479165	T63CP139	GQ479083
T63CP78	GQ479066	T63CP86	GQ479116	T21LP115	GQ479166	T63CP140	GQ479084
T63CP94	GQ479067	T63CP87	GQ479117	T21LP158	GQ479167	T63CP142	GQ479085
T63CP103	GQ479068	T63CP90	GQ479118	T21LP167	GQ479168	T63CP143	GQ479086
T63CP106	GQ479069	T63CP95	GQ479119	T21LP168	GQ479169	T63LP42	GQ479130
T63CP107	GQ479070	T63CP96	GQ479120	T21LP162	GQ479170	T63LP43	GQ479131
T63CP108	GQ479071	T63CP99	GQ479121	T21LP55	GQ479171	T63LP44	GQ479132
T63CP114	GQ479072	T63CP110	GQ479122	T21LP57	GQ479172	T63LP45	GQ479133
T63CP115	GQ479073	T63CP111	GQ479123	T21LP11	GQ479173	T63LP46	GQ479134
T63CP116	GQ479074	T63CP74	GQ479124	T21LP61	GQ479174	T63LP11	GQ479135
T63CP117	GQ479075	T63CP28	GQ479125	T21LP161	GQ479175	T63LP8	GQ479136
T63CP120	GQ479076	T63CP135	GQ479126	T21LP111	GQ479176	T21LP102	GQ479180
T63CP122	GQ479077	T63LP1	GQ479127	T21LP65	GQ479177	T21LP101	GQ479181
T63CP124	GQ479078	T63LP48	GQ479128	T21LP59	GQ479178	T21LP67	GQ479182
T63CP129	GQ479079	T63LP10	GQ479129	T21LP103	GQ479179	T21LP8	GQ479183
T21LP116	GQ479184	T21LP14	GQ479185	T21LP114	GQ479186		

Table S3: Amplification efficiency of different cloned *rhd*□ gene fragment targeted sequences. Sp6/T7 amplicons were purified and 1:10 dilutions were made. The dilutions were used for qPCR as described. Amplified efficiency per cycles was calculated according to the equation:

Efficiency=$10^{\wedge (1/abs(a))}$-1

Cloned gene	Intercept	a	R^2	Amplification efficiency
phnAc (AY540615)	44.31	-4.19	0.982	73.30%
phnAc (GQ479194)	43.78	-3.39	0.993	97.40%
T63CP120 (GQ479076)	39.49	-4.06	0.981	76.20%
T63CP90 (GQ479118)	46.91	-3.51	0.999	92.60%
T63CP49 (GQ479118)	39.11	-3.53	0.998	91.90%

Table S4: Closest hits of the sequences of cloned amplicons from total community DNA of phenanthrene contaminated Luvisol (21 and 63 days) and Cambisol (sampled at Day 63) to *pah-rhdα* gene sequences in GenBank.

Libraries	Number of Sequences	BlastN	Similarity	TblastX	Similarity
T21LP (33)	33	*Burkholderia* sp. Eh1-1 dioxygenase alpha subunit *PhnAc*	98	*Burkholderia* sp. Eh1-1 dioxygenase alpha subunit PhnAc	97
T63LP (37)	32	*Burkholderia* sp. Eh1-1 dioxygenase alpha subunit PhnAc	98	*Burkholderia* sp. Eh1-1 dioxygenase alpha subunit PhnAc	97
	2	*Burkholderia* sp. Eh1-1 dioxygenase alpha subunit PhnAc	98	*Acidovorax* sp. NA3 phenanthrene-degradation gene cluster	100
	2	*Mycobacterium* sp. JLS,rhda	87	*Mycobacterium* sp. JLS	87
	1	*Novosphingobium aromaticivorans* DSM 12444 plasmid pNL1	100	*Novosphingobium aromaticivorans* DSM 12444 plasmid pNL01	100
T63CP(85)	11	*Agrobacterium sanguineum* bphA gene	75	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	77
	9	*Mycobacterium* sp. JLS,rhda	89	*Mycobacterium* sp. JLS	92
	7	*Novosphingobium* sp. H25 isolate 2 ring-hydroxylating dioxygenase	72	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	74
	7	*Novosphingobium aromaticivorans* DSM 12444 plasmid pNL1	83	*Sphingomonas* sp. ZP1 aromatic compound catabolic protein-like gene	81
	6	*Agrobacterium sanguineum* bphA gene	72	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	77
	6	*Novosphingobium aromaticivorans* DSM 12444 plasmid pNL1	84	*Sphingomonas* sp. ZP1 aromatic compound catabolic protein-like gene	81
	6	*Mycobacterium vanbaalenii* PYR-1, *rhda*	98	*Mycobacterium* sp. KMS plasmid pMKMS02	98
	5	*Agrobacterium sanguineum* bphA gene	74	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	77
	5	*Mycobacterium gilvum* PYR-GCK,rhda	98	Mycobacterium gilvum PYR-GCK	100
	4	*Mycobacterium* sp. JLS,rhda	92	*Mycobacterium* sp. JLS	93
	2	*Burkholderia sp.* VUN10,013 dioxygenase alpha subunit gene	99	*Burkholderia* sp. VUN10,013 dioxygenase alpha subunit gene	97
	2	*Novosphingobium* sp. H25 isolate 2 ring-hydroxylating dioxygenase	72	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	78
	2	*Agrobacterium sanguineum* bphA	76	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	78
	2	*Mycobacterium gilvum* PYR-GCK	97	*Mycobacterium gilvum* PYR-GCK	97
	2	*Mycobacterium gilvum* PYR-GCK	95	*Mycobacterium gilvum* PYR-GCK	95
	1	*Mycobacterium gilvum* PYR-GCK,rhda	100	*Mycobacterium gilvum* PYR-GCK	100
	1	*Acidovorax* sp. NA3 phenanthrene-degradation gene cluster	83	*Acidovorax* sp. NA3 phenanthrene-degradation gene cluster	90
	1	*Agrobacterium sanguineum* bphA	74	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	76
	1	*Agrobacterium sanguineum* bphA gene	74	*Novosphingobium* sp. H25 isolate 1 ring-hydroxylating dioxygenase alpha subunit gene	79
	1	*Novosphingobium aromaticivorans* DSM 12444 plasmid pNL1	80	*Novosphingobium aromaticivorans* DSM 12444 plasmid pNL1	80
	1	*Novosphingobium aromaticivorans* DSM 12444 plasmid pNL1	85	*Sphingomonas* sp. ZP1 aromatic compound catabolic protein-like gene	84
	1	*Rhodococcus opacus* naphthalene degradation gene cluster	75	Uncultured bacterium clone NMR-OR05-60 aromatic ring-hydroxylating dioxygenase (*arhD*) gene	55
	1	*Mycobacterium* sp. JLS	86	*Mycobacterium* sp. JLS	94
	1	*Mycobacterium* sp. JLS	93	*Mycobacterium* sp. JLS	90

Sorption experiments. Phenanthrene sorption experiments were carried out in batch incubation experiments according to the guidelines of the OECD (5) with a concentration range of 10 to 50% of the water solubility of phenanthrene (1.1 mg l^{-1}). One gram of soil sample was added to 75 ml matrix solution consisting of 0.01 M $CaCl_2$ and $2 \cdot 10^{-6}$ M $AgNO_3$ spiked with phenanthrene and incubated for 24 hours. The batches were centrifuged and deuterated phenanthrene was added as an internal standard (Ultra Scientific, North Kingstown, RI, USA). Phenanthrene was extracted from the supernatant using solid phase extraction (SPE). 10 ml glass columns were filled with 0.5 g Bakerbond C18 material (J. T. Baker, Philipsburg, NJ, USA) and sequentially conditioned with 5 ml methanol, 5 ml methanol: matrix solution 1:9 and 5 ml matrix solution before extraction of the liquid phase. The columns were then freeze-dried and phenanthrene was eluted with 8 ml of a hexane: dichloromethane 4:1 mixture. The volume of the extract was reduced to 1 ml by rotary evaporation.

Phenanthrene concentration was measured with GC-MS (GC 8000 Top, MS Voyager from Fisons Instruments) with an OPTIMA-5-Accent Fused Silica capillary column (length 30m x inner diameter 25 mm x film thickness 0.25 µm). 1.0 µl of sample was injected splittless at an injector temperature of 280°C. Helium was used as carrier gas with a constant flow rate of 1 ml min^{-1}. The oven program started with 4 min at 85°C, then increased with 15°C min^{-1} to 160°C and stayed at this temperature for 2 min. The temperature was then increased to 230°C at a rate of 5°C min^{-1} and kept there for another 2 min. The concentration of phenanthrene in the solid phase was calculated from the difference between phenanthrene added before incubation and that present in the liquid phase. Pre-test showed that no significant amount of phenanthrene was lost during incubation and thus mass balance could be assumed.

Sorption isotherms were calculated using the solubility-normalized Freundlich isotherm (1) according to the equation: $\log C_{(a)s} = \log K'_f + n^{-1} \log(C_{(a)aq}/S)$ Where $C_{(a)s}$ is the concentration of phenanthrene adsorbed to the solid phase, $C_{(a)aq}$ is the concentration of phenanthrene in solution at equilibrium, S is the super cooled liquid solubility of phenanthrene, n^{-1} is a measure of sorption linearity and K'_f is the Freundlich isotherm coefficient. Three replicates were measured for each phenanthrene concentration level and average concentrations were used for the calculation of the isotherm.

Reference List

1. **Carmo, A. M., L. S. Hundal, and M. L. Thompson.** 2000. Sorption of hydrophobic organic compounds by soil materials: Application of unit equivalent Freundlich coefficients. Environ.Sci.Technol. **34:**4363-4369.
2. **Flocco, C. G., N. C. Gomes, C. W. Mac, and K. Smalla.** 2009. Occurrence and diversity of naphthalene dioxygenase genes in soil microbial communities from the Maritime Antarctic. Environ.Microbiol. **11:**700-714.
3. **Gomes, N. C., L. R. Borges, R. Paranhos, F. N. Pinto, E. Krogerrecklenfort, L. C. Mendonca-Hagler, and K. Smalla.** 2007. Diversity of *ndo* genes in mangrove sediments exposed to different sources of polycyclic aromatic hydrocarbon pollution. Appl.Environ.Microbiol. **73:**7392-7399.
4. **Gomes, N. C., I. A. Kosheleva, W. R. Abraham, and K. Smalla.** 2005. Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. FEMS Microbiol.Ecol. **54:**21-33.
5. **OECD.** 2000. Adsorption/desorption using a batch equilibrium method, OECD guidelines for testing chemicals. OECD publications, Paris.

Chapter 4: Responses of microbial communities in two different soils to phenanthrene spiking as revealed by GeoChip analysis

Guo-Chun Ding[1], Holger Heuer[1], Zhili, He[2], Jianping Xie[2], Jizhong Zhou[2] and Kornelia Smalla[1*]

1Julius Kühn-Institut, Federal Research Centre for Cultivated Plants (JKI), Institute for Epidemiology and Pathogen Diagnostics, Braunschweig, Germany;
2Institute for Environmental Genomics and Department of Botany and Microbiology, University of Oklahoma, Norman, Oklahoma 73072

Abstract

The response of microbes indigenous to two typical central European soils to phenanthrene spiking was investigated by a functional gene array (GeoChip 2.0). A total of 4,192 genes from 11 functional groups involved in several key microbial processes were detected. Twenty-one days after spiking, more genes were detected in samples from contaminated soils than in unspiked soils, with the highest number for polluted Luvisol. A total of 519 genes had significantly higher signal intensities between contaminated and unspiked Luvisol, while very few genes differed between polluted and non-polluted Cambisol soils. Discriminative genes from Luvisol belonging to 11 functional gene groups were detected, which were possibly affiliated to different phyla such as *Proteobacteria, Actinobacteria, Cyanobacteria* and *Firmicutes.* Among the discriminative genes coding for the remediation of organic compounds, most were linked to the degradation of one ring aromatic hydrocarbons, and seven genes were identified as upper pathway genes for polycyclic aromatic hydrocarbon (PAH) degradation. In summary, GeoChip analysis revealed responses of various functional groups affiliated to distant taxa to phenanthrene in Luvisol but not in Cambisol. The differences in soil texture might contribute to the varied responses of soil microbial communities to phenanthrene spiking.

INTRODUCTION

Organic pollutants such as polycyclic aromatic hydrocarbons (PAHs) have attracted a lot of interest to explore their fate in soils due to their potential for bioaccumulation, persistence, transport, and toxicity (11, 14-15). Their effects on soil indigenous microbial communities have been evaluated in many formats such as analysis of the 16S rRNA genes (2, 10, 18), quantitative and diversity analysis of the key functional genes involved in microbial PAH degradation (7-9, 32). These studies nicely pointed out key players in the microbial PAH mineralization network. Several taxonomic groups such as *Sphingomonas, Polaromonas, Burkholderia, Pseudomonas, Mycobacteria, Nocardia,* and *Rhodococcus* were found to be enhanced in PAH degradation in various environmental habitats (4, 10, 23). However, how these organic pollutants influence soil key functions such as carbon, nitrogen, and phosphorus cycling has hardly been evaluated comprehensively, mainly due to the lack of comprehensive tools. The development of microarrays now allows the simultaneous evaluation of thousands of functional genes. The functional gene array GeoChip 2.0 developed by He et al. 2007 (12) consists of 24,243 oligonucleotide probes targeting over 10,000 genes in more than 150 functional groups involved in nitrogen, carbon, sulfur and phosphorus cycling, metal reduction and resistance, and organic contaminant degradation (12). It has been successfully applied to analyze the diversity and structure of functional genes in different habitats, such as water (12, 16, 27), soil (20), sediment (31), extreme environments (28, 30), and oil-contaminated sites (21, 22). All these studies have provided ecological insights into our understanding of relationships between microbial communities and environmental factors, such as the influence of physicochemical properties (e.g., pH, temperature) on microbial processes, and changes of microbial processes during bioremediation of contaminants (e.g., heavy metals, organic pollutants).

Soil samples from a microcosm experiment with a silt loam (Luvisol) and a clay loam (Eutric Cambisol) were analyzed previously by PCR targeting *PAH-RHDα* genes which are key genes for microbial mineralization of PAHs. While in control soils no *PAH-RHDα* genes were detected, phenanthrene spiking resulted in an enrichment of populations carrying *PAH-RHDα* genes. The abundance and diversity of *PAH-RHDα* genes was soil type dependent. However, it remained unknown whether phenanthrene spiking altered other soil functional groups as well, such as

those involved in carbon, nitrogen, phosphate cycling. Therefore, in this study GeoChip 2.0 was used to explore shifts in the functional genes in two soils in response to phenanthrene spiking.

MATERIALS AND METHODS

Experimental design and sampling. The set-up of the experiment was described previously (7). Briefly, soils were contaminated with phenanthrene to reach a final concentration of 2 mg g^{-1} soil. Four replicate microcosms of Luvisol or Cambisol, either phenanthrene-treated or not, were incubated at room temperature (23°C) in the dark. Samples were taken at days 0, 21 and 63 and kept at -20°C before DNA extraction with Bio101 DNA spin kit for soil (Qbiogene, Heidelberg, Germany). Soil parameters, phenanthrene analysis, quantitative real-time PCR (qPCR) for 16S rRNA genes, studies on the abundance and diversity of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase (*PAH-RHDα*) genes as well as the soils' sorption ability on phenanthrene were described previously (7). Samples collected at days 0 and 21. Total community DNA (TC-DNA) was extracted by means of Bio-101 DNA spin kit for soil (Qbiogene, Heidelberg, Germany). The following abbreviations are used in the text for the treatments: T0C, Cambisol soil at day 0; T0L, Luvisol soil at day 0; T21CP, phenanthrene-spiked Cambisol at day 21; T21CA, Cambisol control at day 21; T21LP, phenanthrene-spiked Luvisol at day 21; T21LA, Luvisol control at day 21.

DGGE analysis of 16S rRNA gene amplicons. PCR amplification of 16S rRNA gene fragments from TC-DNA and denaturing gradient gel electrophoresis (DGGE) of the amplicons was done as previously described (10, 13). Software package GelCompare 4.5 was used for cluster analysis of DGGE profiles. Dendrograms were constructed based on pair wise Pearson correlations and the unweighted pair group method using arithmetic averages (UPGMA). The pair wise Pearson correlation matrix was subjected to a permutation test for significant differences between treatments using the PERMTEST software developed by Kropf et al. (17).

GeoChip analysis. GeoChip hybridization was performed as follows: total community DNA was subjected to whole community genome amplification and then

labeled with Cy-5 (29). Labeled products were purified and subjected to hybridization (29). The scanned images were gridded and quantified with software package ImaGene 6.0 (Biodiscovery Inc., EL Segundo, CA, USA). Noise data were removed according to He et al (12). When all corresponding probes for a particular gene were detected this gene was reported as being present. The hybridization score for each gene was the average signal intensity of all corresponding probes. Three individual hybridizations for each of the four replicates were done. But due to the technical problems and quality control procedures, only 46 slides for 21 samples (Table 1) were suitable for further analysis. All data from technical replicates were combined to their corresponding biological replicates using the average hybridization scores for detected genes.

Statistical analysis. Multiple one-way ANOVA in conjunction with TukeyHSD tests were used to identify genes with significant differences in signal intensity between polluted and non-polluted soils and between soil types. The statistical analysis and data summaries were performed with software package R 2.11. Kruskal's nonmetric multidimensional scaling analyses was done with R packages MASS based binary distance between different samples.

Analysis of genes involved in metabolic pathways for xenobiotics degradation. The role of significantly responding genes in biodegradation pathways of xenobiotics was studied by the following two approaches. Firstly, all amino acid sequences encoded by significantly enriched genes were downloaded from NCBI and analyzed by BLAST-P against K genes on Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/). Based on the BLAST-P report, KEGG gene orthology identity numbers were retrieved and mapped against the pathway of microbial metabolism in diverse environments (http://www.kegg.jp/kegg-bin/show_pathway?map01120) (24). Secondly, genes in the upper pathway for PAHs mineralization were retrieved based on BLAST-N reports of known *PAH-RHDα* genes. All closest hits were downloaded and amino acid sequences for corresponding genes in the upper pathway were selected and checked manually. A local database was constructed with amino acid sequences coded by 218 genes in known upper pathways (*nah, phn, ndo, bph, nar, pdo, dox, nid*) from 40 sequences (accession number in supplement) in NCBI. The retrieved amino sequences for

significantly responding organic remediation genes were analyzed by BLAST-P against the local database using standalone BLAST tools. Those genes with significant hits (1 e-30 and >75% conserved) in the database were regarded as upper pathway genes.

RESULT

Bacterial community shifts after phenanthrene spiking detected by PCR-DGGE. To compare the responses of the bacterial communities to phenanthrene spiking in both soils, PCR amplicons of 16S rRNA genes were analyzed by DGGE. Bands with strongly increased intensity (Fig. 1a: a and b) were observed in polluted Luvisol while in the polluted Cambisol only bands with slightly enhanced intensity could be seen (Fig. 1a, c). Permutation tests confirmed a significant effect of phenanthrene contamination on bacterial communities in both soils, when spiked and control soils were compared 21 days after treatment ($P<0.05$).Stronger responses to phenanthrene spiking were detected for the Luvisol. The bacterial community profiles shared an average of only 27.2% similarities between the polluted Luvisol and the corresponding control, while 54.3% similarity was shared between polluted Cambisol and its control. The DGGE fingerprints of the bacterial communities of the two soil types differed significantly ($P<0.0001$). Interestingly, the bacterial profiles of spiked treatments were more similar among the two soil types than the untreated controls of the two soils, indicating that bacterial community structures were shaped by phenanthrene spiking in a similar way among replicates.

Effects of phenanthrene spiking on functional gene abundance. To characterize the overall changes of functional genes in both soils to phenanthrene, the functional gene array (GeoChip 2.0) was used to compare the microbial communities of spiked and untreated soils. A total of 4,192 genes in 11 functional gene categories were detected from both soils. Most detected genes were related to organic pollutant remediation (ORGs), heavy metal resistance (MIT), nitrogen cycling or carbon cycling. More genes were detected in samples from the phenanthrene contaminated soils with the highest number for polluted Luvisol compared to the corresponding control soil or the soil taken at day 0 (Table 1). More genes were detected that might be due to the enrichment of certain populations in the polluted soils.

To identify genes responding to phenanthrene or soil type, multiple one-way ANOVA in conjunction with TukeyHSD tests were used ($P<0.01$). More than 12% of the detected genes (4,192) were increased in phenanthrene-spiked Luvisol compared to the untreated control (Table 2). In addition to genes involved in organic pollutants remediation, genes in other categories such as nitrogen cycling, heavy metal resistance, carbon cycling, or sulfate reduction were enriched in the polluted Luvisol. For most of these functional groups, the ratio of significantly enriched genes varied between 6 to 16% of the genes detected for the corresponding functional group. A high percentage of responding genes was found in the PEP (phosphoenolpyruvate) group, where two out of three genes were of significantly higher signal intensity than the unpolluted control. This was followed by methane oxidation genes where 16.4% of the detected genes were enriched. Only ca. 11% detected organic remediation genes were enriched although in this gene category the highest number of responding genes was found. The lowest percentage of responding genes was found for gene categories carbon fixation (CFIX) only of 6%. Much less responding genes were detected for the Cambisol (0.2%). No clear differences were observed between the soil types at both sampling times.

Overall functional gene pattern. To study the patterns for all functional genes and organic remediation genes, nonmetric multidimentional scaling was performed based on binary distance between samples (Fig. 2 a and b). All polluted soil samples except one replicates from T21CP were grouped together indicating that similar functional genes might respond to phenanthrene spiking in both soils. Soil samples at day 0 and replicates of T21LA separated from each other while the replicates of T21CA was more close to each other suggested an incubation effect on the functional gene structure for the Cambisol.

Functional genes increased by phenanthrene spiking.Only eight genes with significantly higher signal intensity were observed in phenanthrene-contaminated Cambisol compared to the non-contaminated control. Five of them (NCBI gi: 40890267, 40890329, 14289342, 24575091, 26991304) were related to organic pollutant remediation but none of the genes enriched were from the upper pathway of PAHs mineralization. Two genes with significantly enhanced signal intensity

(3298347, 19570948) were related to the carbon cycle and one (401315) related with the nitrogen cycle.

A total of 519 genes were enriched in the polluted Luvisol and 158 of these were related to organic pollutant remediation. To analyze the enriched genes in the microbial metabolism pathways, 498 amino acid sequences (gene identity number in supplement) from NCBI were compared by BLAST-P to genes in the KEGG database. Altogether 214 genes enriched in the polluted Luvisol which corresponded to 66 KEGG gene orthology identity numbers were mapped in reference pathways (Fig. S1). Among those mapped genes, 87 belonged to ORGs of which most were related with the degradation pathway for one ring aromatic hydrocarbons such as benzoate (29), nitrotoluene (28), atrazine (22), aminobenzoate (18), styrene (15), toluene (14), xylene (12), fluorobenzoate (12), dioxin (9) or chlorohydrocarbons such as chlorocyclohexane (26) or chloroalkane (24). Also 13 genes were linked with the mineralization of PAHs such as naphthalene, pyrene, 94luorine, benz[a]anthracene, and benzo[a]pyrene, but only one (*rhda*,26080263) from *Mycobacterium gilvum* could be identified as an upper pathway gene. This result suggested that phenanthrene spiking enhanced the groups degrading low molecular weight aromatic hydrocarbons.

Microbial PAHs degradation is initiated by enzymes encoded by the upper pathway genes. To identify the involvement of responding organic pollutant remediation genes in the upper pathway of PAHs degradation, all amino acid sequences encoded by enriched genes in polluted Luvisol were compared to known upper pathway genes by BLAST-P. Six genes (*nahAa*: 1255667, *nahAc*: 2828018, 3170519, *nahQ*: 6226946 and *narR*-like: 38524451, 38524452) were identified as upper pathway genes for PAHs degradation.

Seventy-one out of 214 mapped genes were involved in nitrogen cycling. An increase of nitrogen fixing bacteria in phenanthrene-spiked Luvisol was suggested by the enrichment of seventeen *nifH* genes. Bacteria involved in denitrification might be also enhanced because of enriched genes, such as *narG* (11)*, nirS* (8)*, norB* (4) and *nirK* (2). All mapped nitrification genes (21) coded for urease.

The taxonomy of corresponding organisms which harbor significantly enriched genes in the polluted Luvisol were retrieved from NCBI taxonomy and summarized in Table 3. Altogether 367 genes could be linked to *Bacteria* and *Archaea*. The majority of the genes (360) could be affiliated with *Proteobacteria, Actinobacteria, Cyanobacteria* and *Firmicutes*. Most ORGs and METs were affiliated with *Grammaproteobacteria, Alphaproteobacteria* and *Betaproteobacteria*. The majority (62.7%) of NIT genes was affiliated with *Proteobacteria* and also 14.7% with *Cyanobacteria*. 50% DSR genes were affiliated with *Deltaproteobacteria*. Overall, the increased signal intensity of functional genes from various different phyla indicated that population belonging to distantly related taxa responded to the phenanthrene spiking.

Enrichment of genes belonging to several functional groups might due to some of these genes exist in the same microbes. Therefore, different functional groups were clustered based on pair-wise Pearson correlation indices which were calculated based on the number of significant responding genes at different genus (Table S1). Interestingly, ORG and MET share highest similarity of 62.1%, followed by 49.1% between NIT and ORGs and 45.0% between NIT and MET. The result showed that a considerable part of the genes belonging to ORG, METs and NITs come from hosts belonging to the same genus. Therefore, it is possible that enriched genes belonging to the MET and NIT gene categories coexisted with ORG genes in the same population. The remaining gene categories shared less than 12.4% similarity (Fig. 3).

DISCUSSION

The effects of polycyclic aromatic hydrocarbons on the soil microbial community are of great interest due to their harmful effects on the environment and human health (11, 14-15). Important taxonomic groups and key functional genes were pinpointed (11, 14-15). By GeoChip analysis, this study explored the global changes of functional genes in two soils after a 21-day phenanthrene spiking.

In our study, a higher number of functional genes were detected in samples from contaminated soils compared to untreated soils. This should not be interpreted as an increased functional diversity after phenanthrene spiking but the increased number of

detected genes is most likely due to an increased relative abundance of certain populations. The functional genes of these enriched populations passed the detection threshold of the GeoChip analysis. This is supported by at least two lines of evidence. Firstly, the increased abundance of some populations in spiked Luvisol was evidenced in the DGGE fingerprints by bands of increased intensity for both phenanthrene-spiked soils, with the strongest bands in the polluted Luvisol. Secondly, at this sampling time *PAH-RHDα* genes could only be detected in the contaminated Luvisol but not in the Cambisol (7).

Interestingly, the enriched genes could be linked to distant taxonomic groups instead of only few known PAH mineralization taxa, and to various biochemical functions such as heavy metal resistance or nitrogen cycling which have little to do with organic pollutant remediation. For example, sulfite reduction genes such as *dsrA, dsrB* which were frequently found in *Desulfobacterales*, urease genes affiliated with a genus of *Cyanobacteria* and heavy metal resistance genes associated with *Firmicutes* were all enriched in the polluted Luvisol. The spiking of phenanthrene probably not only enhanced organic pollutant remediation group, but also other functional groups involved in carbon, nitrogen, and sulfur cycling. Most enriched ORG, MET and NIT genes detected were related with known PAH mineralization taxa such as *Burkholderiales, Pseudomonas, Enterobacteriales*, *Actinomycetales* and *Alphaproteobacteria*. Correlation analysis between different functional groups suggested that a part of enriched genes related with metal resistance or nitrification might originate from the same host as enriched organic remediation genes. For other functional groups such as carbon degradations, denitrification, sulfate reduction, co-metabolism might be a reason to explain their enrichment. Nitrite is an electron acceptor under anaerobic conditions. The increased abundance of nitrite reduction genes in the polluted Luvisol suggested that bacterial populations involved in denitrification were increased possibly in micro-sites with depleted oxygen level. Most enriched genes belonging to ORGs are linked with the degradation of one-ring aromatic hydrocarbons and chloro-hydrocarbons, with few genes in the upper pathway for polycyclic aromatic hydrocarbon degradation. Even though microbes harboring an upper pathway for PAH degradation are essential for the initial steps in PAHs degradation, microbes with lower pathways seemed to profit most from phenanthrene mineralization in the polluted Luvisol. But neither the number of genes

nor total signal intensities detected for different taxa or gene categories were expected to reflect the structural and functional community composition because only genes represented on the chip could be detected, and the number of representative genes matters.

The GeoChip results suggested that the upper pathway genes *nahAc* and *narAc* were enriched in the polluted Luvisol. These *nah* genes previously often reported to be located on IncP-9 plasmids from *Pseudomonas* (19, 26). The *nar* genes were typically found in strains of the genus *Rhodococcus* (1, 6). But according to PCR-based PAH-RHDα gene analysis, only *phnAc* gene was found to be enriched (7). Similar *phnAc* genes were detected by GeoChip hybridization with equal signal intensities for polluted and control soils. Most likely, the relative abundance of *phnAc* genes in polluted Luvisol might be still very low which did not allow them to be reliably detected by GeoChip. On the other side, even though a lot of efforts were invested to broaden the spectrum of the applied PCR system for *PAH-RHDα* genes, and *in silico* analysis suggests that both *nahAc* and *narA* genes could be amplified, it is still possible that previously undescribed homologous genes in the soil samples could not be targeted by the primers. Generally, microarrays were supposed to detect known genes which were represented on the GeoChip 2.0 but it is also possible that previously undescribed homologous genes could hybridize with the probes and therefore were misjudged as known genes. The tool was developed based on known gene sequences then, but so far the ecological diversity of many genes is still poorly uncovered.

Even though more genes were detected from phenanthrene-spiked Cambisol, very few genes were of significantly higher signal intensities. This result suggested that the response of microbial community in the Cambisol to phenanthrene spiking was very weak or that diverse populations in the Cambisol enriched but the elevated levels of most enriched groups were too low. Previous chemical studies show that a similar decreased amount of phenanthrene was observed for both soils at this sampling time (7). Therefore, it is assumed that many taxa enriched in the Cambisol. The difference between soil physicochemical characters could influence the responses of microbial communities to phenanthrene pollution. High contents of silt- and clay-sized particles could reduce the soil pore connectivity thus increasing

diversity of the soil bacteria (3). The Cambisol has much higher clay content than the Luvisol. Sorption and low pore connectivity in the Cambisol could have contributed to the many microbial groups to phenanthrene spiking. Despite the soil type dependent microbial response to phenanthrene spiking, it is also worth to note that a relative high overlap of functional genes shared by most polluted samples of both soil types.

Contrasting bacterial community structures were revealed by PCR-DGGE analysis for both soil types. However, the GeoChip analysis revealed only very few discriminative genes between the soil types. This might lead to the conclusion that a similar functional gene structure was shared by both soils with a significantly different bacterial community structure. However, only limited numbers of genes were detected for uncontaminated soil samples, and the majority of the functional genes were still unknown. Soil microbial communities are extremely diverse (5, 25). Thus for most microbes their relative abundance was too low to detect their functional genes reliably by GeoChip 2.0. Hence, sensitivity of the tools should be considered seriously when the functional diversity/potential is studied for soils or other complex environmental samples.

In summary, the GeoChip analysis confirmed the strong responses of the microbial soil community to phenanthrene spiking in Luvisol, suggesting that various functional groups linked with distant taxonomic groups were involved in the mineralization network for PAHs.

ACKNOWLEDGMENTS

This work was supported by DFG SPP1315 (SM59/8–1), U.S. Department of Energy through the Environmental Remediation Science Program (ERSP) and a Scientific Focus Area Program, ENGIMA (DE-AC02-05CH11231).

REFERENCES

1. **Andreoni, V., S. Bernasconi, M. Colombo, J. B. van Beilen, and L. Cavalca.** 2000. Detection of genes for alkane and naphthalene catabolism in Rhodococcus sp. Strain 1BN. Environ. Microbiol. **2:**572-577.
2. **Bordenave, S., M. S. Goni-Urriza, P. Caumette, and R. Duran.** 2007. Effects of heavy fuel oil on the bacterial community structure of a pristine microbial mat. Appl. Environ. Microbiol. **73:**6089-6097.
3. **Carson, J. K., V. Gonzalez-Quinones, D. V. Murphy, C. Hinz, J. A. Shaw, and D. B. Gleeson.** 2010. Low pore connectivity increases bacterial diversity in soil. Appl. Environ. Microbiol. **76:**3936-3942.
4. **Chain, P. S., V. J. Denef, K. T. Konstantinidis, L. M. Vergez, L. Agullo, V. L. Reyes, L. Hauser, M. Cordova, L. Gomez, M. Gonzalez, M. Land, V. Lao, F. Larimer, J. J. LiPuma, E. Mahenthiralingam, S. A. Malfatti, C. J. Marx, J. J. Parnell, A. Ramette, P. Richardson, M. Seeger, D. Smith, T. Spilker, W. J. Sul, T. V. Tsoi, L. E. Ulrich, I. B. Zhulin, and J. M. Tiedje.** 2006. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. PNAS **103:**15280-15287.
5. **Curtis, T. P., W. T. Sloan, and J. W. Scannell.** 2002. Estimating prokaryotic diversity and its limits. PNAS **99:**10494-10499.
6. **Di Gennaro, P., P. Terreni, G. Masi, S. Botti, F. De Ferra, and G. Bestetti.** 2010. Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. Appl. Microbiol. Biotechnol. **87:**297-308.
7. **Ding, G. C., H. Heuer, S. Zühlke, M. Spiteller, G. J. Pronk, K. Heister, I. Kogel-Knabner, and K. Smalla.** 2010. Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes by using a novel PCR detection system. Appl. Environ. Microbiol. **76:**4765-4771.
8. **Flocco, C. G., N. C. Gomes, C. W. Mac, and K. Smalla.** 2009. Occurrence and diversity of naphthalene dioxygenase genes in soil microbial communities from the Maritime Antarctic. Environ. Microbiol. **11:**700-714.
9. **Gomes, N. C., L. R. Borges, R. Paranhos, F. N. Pinto, E. Krögerrecklenfort, L. C. Mendonça-Hagler, and K. Smalla.** 2007. Diversity of *ndo* genes in mangrove sediments exposed to different sources of

polycyclic aromatic hydrocarbon pollution. Appl. Environ. Microbiol. **73:**7392-7399.

10. **Gomes, N. C., I. A. Kosheleva, W. R. Abraham, and K. Smalla.** 2005. Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. FEMS Microbiol. Ecol. **54:**21-33.
11. **Habe, H., and T. Omori.** 2003. Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. Biosci. Biotech. Biochem. **67:**225-243.
12. **He, Z., T. J. Gentry, C. W. Schadt, L. Wu, J. Liebich, S. C. Chong, Z. Huang, W. Wu, B. Gu, P. Jardine, C. Criddle, and J. Zhou.** 2007. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J **1:**67-77.
13. **Heuer, H., M. Krsek, P. Baker, K. Smalla, and E. M. Wellington.** 1997. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl. Environ. Microbiol. **63:**3233-3241.
14. **Johnsen, A. R., L. Y. Wick, and H. Harms.** 2005. Principles of microbial PAH-degradation in soil. Environ. Pollut. **133:**71-84.
15. **Kanaly, R. A., and S. Harayama.** 2000. Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. J. Bacteriol. **182:**2059-2067.
16. **Kimes, N. E., J. D. Van Nostrand, E. Weil, J. Zhou, and P. J. Morris.** 2010. Microbial functional structure of *Montastraea faveolata*, an important Caribbean reef-building coral, differs between healthy and yellow-band diseased colonies. Environ. Microbiol. **12:**541-556.
17. **Kropf, S., H. Heuer, M. Grüning, and K. Smalla.** 2004. Significance test for comparing complex microbial community fingerprints using pairwise similarity measures. J. Microbiol. Methods **57:**187-195.
18. **Kumar, M., and S. Khanna.** 2010. Diversity of 16S rRNA and dioxygenase genes detected in coal-tar-contaminated site undergoing active bioremediation. J. Appl. Microbiol. **108:**1252-1262.
19. **Lambertsen, L. M., S. Molin, N. Kroer, and C. M. Thomas.** 2004. Transcriptional regulation of pWW0 transfer genes in *Pseudomonas putida* KT2440. Plasmid **52:**169-181.

20. **Leigh, M. B., V. H. Pellizari, O. Uhlik, R. Sutka, J. Rodrigues, N. E. Ostrom, J. Zhou, and J. M. Tiedje.** 2007. Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). ISME J **1:**134-148.

21. **Liang, Y., G. Li, J. D. Van Nostrand, Z. He, L. Wu, Y. Deng, X. Zhang, and J. Zhou.** 2009. Microarray-based analysis of microbial functional diversity along an oil contamination gradient in oil field. FEMS Microbiol. Ecol. **70:**324-333.

22. **Liang, Y., J. D. Van Nostrand, Y. Deng, Z. He, L. Wu, X. Zhang, G. Li, and J. Zhou.** 2011. Functional gene diversity of soil microbial communities from five oil-contaminated fields in China. ISME J **5:**403-413.

23. **Mattes, T. E., A. K. Alexander, P. M. Richardson, A. C. Munk, C. S. Han, P. Stothard, and N. V. Coleman.** 2008. The genome of Polaromonas sp. Strain JS666: insights into the evolution of a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to biotechnology. Appl. Environ. Microbiol. **74:**6405-6416.

24. **Okuda, S., T. Yamada, M. Hamajima, M. Itoh, T. Katayama, P. Bork, S. Goto, and M. Kanehisa.** 2008. KEGG Atlas mapping for global analysis of metabolic pathways. Nucleic Acids Res. **36:**W423-426.

25. **Schloss, P. D., and J. Handelsman.** 2006. Toward a census of bacteria in soil. PloS Comput. Biol. **2:**e92.

26. **Sota, M., H. Yano, A. Ono, R. Miyazaki, H. Ishii, H. Genka, E. M. Top, and M. Tsuda.** 2006. Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase. J. Bacteriol. **188:**4057-4067.

27. **Van Nostrand, J. D., W. M. Wu, L. Wu, Y. Deng, J. Carley, S. Carroll, Z. He, B. Gu, J. Luo, C. S. Criddle, D. B. Watson, P. M. Jardine, T. L. Marsh, J. M. Tiedje, T. C. Hazen, and J. Zhou.** 2009. GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer. Environ. Microbiol. **11:**2611-2626.

28. **Wang, F., H. Zhou, J. Meng, X. Peng, L. Jiang, P. Sun, C. Zhang, J. D. Van Nostrand, Y. Deng, Z. He, L. Wu, J. Zhou, and X. Xiao.** 2009. GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de Fuca Ridge hydrothermal vent. PNAS **106:**4840-4845.

29. **Wu, L., X. Liu, C. W. Schadt, and J. Zhou.** 2006. Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl Environ. Microbiol. **72:**4931-4941.

30. **Xie, J., Z. He, X. Liu, J. D. Van Nostrand, Y. Deng, L. Wu, J. Zhou, and G. Qiu.** 2011. GeoChip-based analysis of the functional gene diversity and metabolic potential of microbial communities in acid mine drainage. Appl. Environ. Microbiol. **77:**991-999.

31. **Xu, M., W. M. Wu, L. Wu, Z. He, J. D. Van Nostrand, Y. Deng, J. Luo, J. Carley, M. Ginder-Vogel, T. J. Gentry, B. Gu, D. Watson, P. M. Jardine, T. L. Marsh, J. M. Tiedje, T. Hazen, C. S. Criddle, and J. Zhou.** 2010. Responses of microbial community functional structures to pilot-scale uranium *in situ* bioremediation. ISME J **4:**1060-1070.

32. **Yagi, J. M., and E. L. Madsen.** 2009. Diversity, abundance, and consistency of microbial oxygenase expression and biodegradation in a shallow contaminated aquifer. Appl. Environ. Microbiol. **75:**6478-6487.

TABLE 2. Average of genes detected for different samples

	T0C (2)	T21CA(4)	T21CP(4)	T0L(4)	T21LA(4)	T21LP(3)
ORG	66±28.3	271±102.8	472±249.8	68±59.4	106±52.8	678±237.7
MET	34±12	136±39.5	236±121.7	34±33.7	56±22.4	346±116
NIT	36±12.7	129±32.5	188±85.1	32±27	50±20.7	264±72
NRED	23±2.8	96±25.1	147±71	18±20	29±14.1	194±78.8
CDEG	14±6.4	62±24.9	112±63	12±10	20±12	180±59.6
DSR	18±7.1	68±23.5	108±54.1	16±12.8	26±11.1	153±48.8
CFIX	12±1.4	36±8.5	54±29.3	12±8.2	20±10.1	74±26.1
NFIX	7±2.8	25±10	39±21.6	6±6.2	11±5.1	56±19
Methane gen	3±1.4	15±5.3	27±12.7	4±5.2	8±4.8	38±18
Methane ox	8±3.5	22±3.6	34±15	7±5.2	10±4.2	44±7.6
PER	0±0	0±0	1±0.8	0±0	0±0	2±0.6
total	220±78.5	858±264.8	1417±719.6	210±186.3	334±146.5	2028±679

Note: (number of replicates analyzed); average of genes detected ± standard deviation; ORG: organic remediation genes, MET: metal resistance genes, NIT: nitrificiation; NRED: denitrification; CDEG: carbon degradation; DSR: Sulfite reduction. CFIX: carbon fixation; NFIX: nitrogen fixation; Methane gen: methane production; Methane ox: methane oxidation; PEP: phosphoenolpyruvate

TABLE 3. Number of genes with significant difference in signal intensity ($p<0.01$)

	T0L- T0C	T21CP- T21CA	T21LA- T21CA	T21LP- T21LC	Total
ORG	0	5	1	158	1396
MET	0	0	0	104	685
NIT	1	1	3	75	499
NRED	0	0	0	47	445
CDEG	0	2	1	49	377
DSR	0	0	0	35	301
CFIX	0	0	0	11	179
NFIX	0	0	0	18	132
methanegen	0	0	0	8	96
methaneox	0	0	0	12	73
PER	0	0	0	2	3
Total	1	8	5	519	4192

Table 4: Number of discriminative genes to phenanthrene spiking in Luvisol and their taxonomic affiliation

domain	phylum	class	order	family	ORG	MET	NIT	CDEG	NRED	DSR	other	Total
Bacteria	*Proteobacteria*	*Alphaproteobacteria*			27	10	14	1	2	0	4	58
		Betaproteobacteria	*Burkholderiales*		28	9	4	1	2	0	3	47
			other		1	1	7	0	2	0	0	11
		Gammaproteobacteria	*Pseudomonadales*	*Pseudomonadaceae*	31	11	6	1	0	0	0	49
				Enterobacteriales	10	18	6	2	1	0	0	37
				other	11	12	0	5	0	0	2	30
		other			1	10	2	0	0	14	0	27
	Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Nocardiaceae*	9	0	0	0	0	0	0	9
				Mycobacteriaceae	9	6	5	1	0	0	0	21
				other	9	5	3	3	0	0	0	20
			Bifidobacteriales	*Bifidobacteriaceae*	0	1	0	0	0	0	0	1
	Cyanobacteria				3	2	11	1	1	0	3	21
	Firmicutes				0	9	4	3	0	0	2	18
	other				3	4	0	3	0	1	0	11
Archaea					2	0	1	0	0	0	4	7
others					14	5	12	28	27	13	32	131
Total					158	103	75	49	35	28	50	498

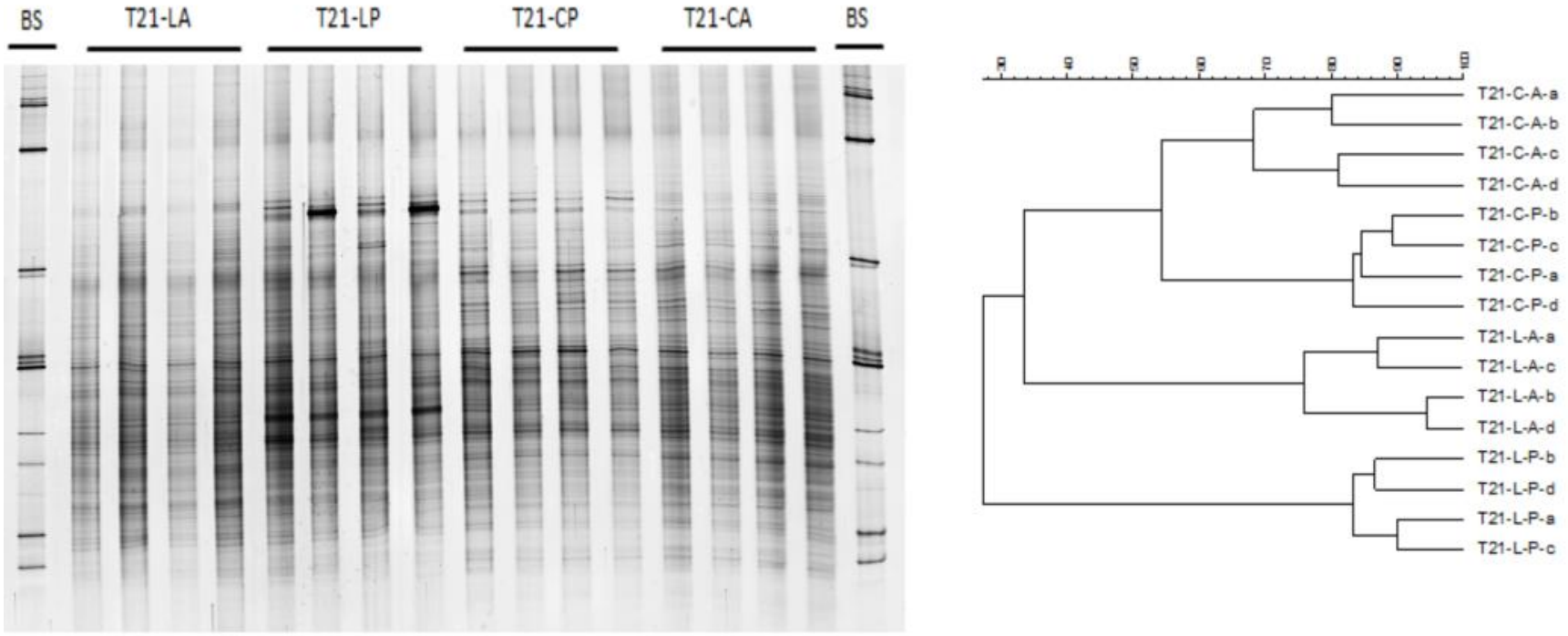

Fig. 1: Bacterial DGGE fingerprint (a) and the corresponding UPGMA clusters (b) for soil spiked with phenanthrene (LP, CP) or not (LA, CA) at day 21. BS : Bacterial standard consisting of 16S rRNA gene fragment from 11 bacterial isolates.

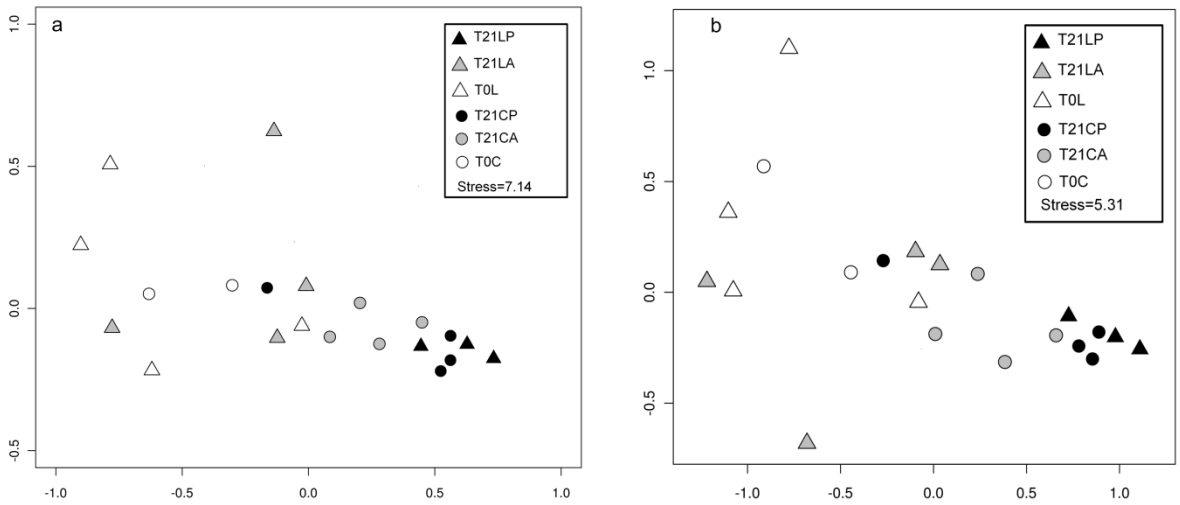

Fig. 2: Nonmetric multidimentional scaling using all functional genes(a) and organic remediation genes (b) detected for soils (L,C) at day 0 (T0) and soils spiked with phenanthrene(LP,CP) or not (LA,CA) at day 21(T21).

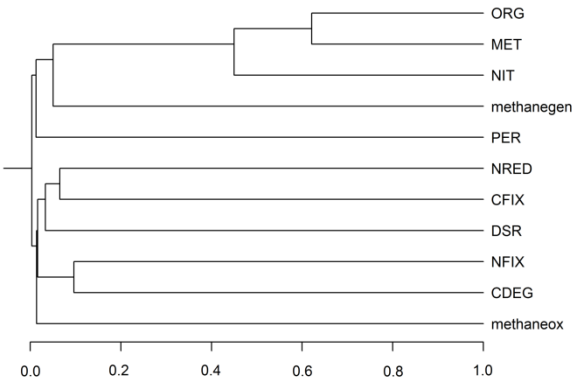

Fig. 3: Furthest-Neighbor-joining cluster of different gene groups based on Pearson correlation distance using numbers of enriched genes belonging to different bacterial genus.

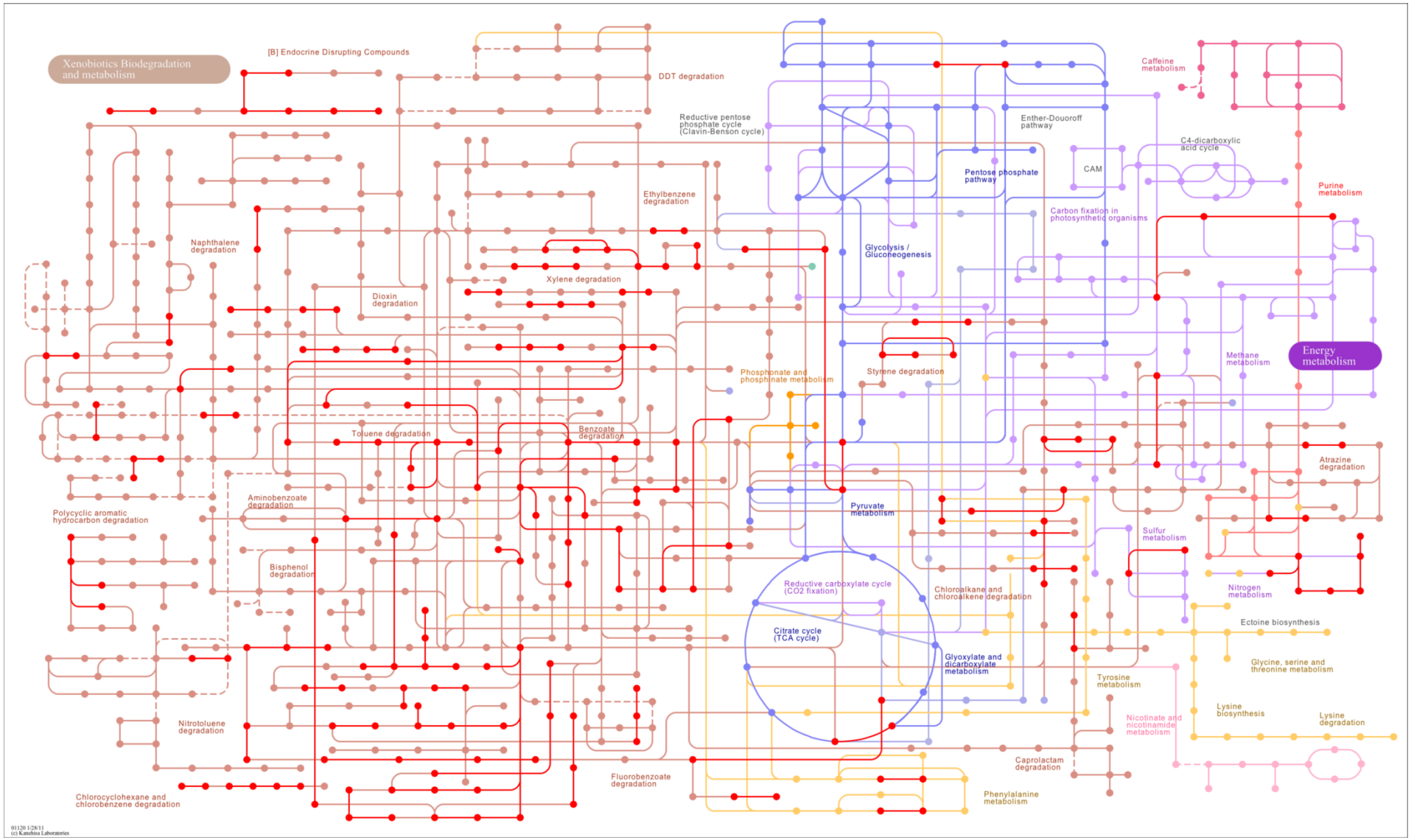

Fig. s1 Metabolic pathway for genes (dark lines) with significantly higher signal intensity in polluted in luvisol.

Chapter 5: Establishment of a novel primer system targeting conserved stretches in the *rep-oriV* region of IncP-9 plasmids

Guo-Chun Ding[1], Holger Heuer[1], Christopher Thomas[2] and Kornelia Smalla[1]

Julius Kühn-Institut (JKI) – Federal Research Centre for Cultivated Plants, Messeweg 11-12, 38104 Braunschweig, Germany

Abstract

IncP-9 plasmids are known to often carry PAH degradative genes and due to their host range they are assumed to contribute to the adaptation of *Pseudomonas* species to pollution. A novel primer system targeting conserved stretches in the *rep-oriV* region was developed, validated and applied to study the response of bacteria harboring IncP-9 plasmids to phenanthrene in two soils from a microcosm experiment.

Introduction

Plasmids are important mobile gene elements which contribute to the adaption of *Bacteria* to changing environments. By conjugative transfer, IncP-9 plasmids can shuttle among a wide spectrum of *Pseudomonas* species [1]. Constitutive expression of genes for mating pair formation suggested that these plasmids are ready for transfer [2]. But expression is low and frequently negatively regulated to reduce the metabolic burden on the host [2]. Beyond these genes for transfer, IncP-9 plasmids carry genes and characteristic regions for autonomous replication and stable inheritance in their natural host [3]. Comparison of the whole plasmid genome of NAH7, pDTG1 and pWWO suggested IncP-9 plasmids have a ~25 kb region which supports the essential functions such as replication, transfer and stable maintenance [1]. Many IncP-9 plasmids were found to enable their host to utilize aromatic compounds such as naphthalene, phenanthrene. Three out of four completely sequenced plasmids NAH7, pDTG1 and pNAH20 bear the operon for the upper pathway of naphthalene degradation which code for enzymes catalyzing the transformation of naphthalene to salicylate under aerobic condition [4,5]. In addition, genes conferring mercury resistance, UV resistance, antibiotic resistance were found on some of the IncP-9 plasmids analyzed [6,7]. The location of accessory genes on composite transposons not only provides the chance to integrate into the host chromosome but also to supply the opportunity for quick adaption by recruiting new genes [4,6]. Isolates carrying IncP-9 plasmids were obtained from Asia, Europe and North America indicating an extremely wide geographic distribution [1,8]. Plasmids belonging to the IncP-9 group were isolated from various environmental niches such as coal tar-contaminated soil, oil-contaminated soil, and manure exposure to antibiotics [9,10]. In contrast to previous views, it is assumed that incP-9 plasmids

contribute globally to the adaption of *Pseudomonas* species to polluted environments by aromatic compounds, heavy metals and antibiotics.

So far most of the knowledge on IncP-9 plasmid diversity is based on isolates and thus our insights might be biased due to the limitations of cultivation-dependent methods. Using specific primers targeting at genes on the backbone provide a chance to study the prevalence and diversity of IncP-9 plasmids [1,8,9,11]. Except for choosing target genes which play an essential role for the survival of the plasmids, enough information on their diversity is also important. The *rep-oriV* region is essential for autonomous replication of IncP-9 plasmids [12,13]. The Rep protein can bind to *oriV* regions and allow further interaction with replication machinery of the host [13]. Sevastsyanovich *et al.* (2008) showed that 1.3 kb *rep-oriV* was sufficient for the autonomous replication of IncP-9 plasmid in their *Pseudomonas* host [1]. Interestingly, overexpression of the *rep* gene could result in its replication in *E. coli* [13]. Specific primers were designed, applied to study environmental samples [9,11]. Krasowiak et al. (2002) found that the IncP-9 like plasmids were prevalent in various environments by a specific primer targeting at *rep* gene [9]. But only very limited numbers of sequences were available at the time and the diversity of rep genes for IncP-9 plasmids had been poorly understood. Recently, Sevastsyanovich et al. (2008) analyzed their *rep* gene and *oriV* sequences of IncP-9 plasmids collected all over the world. The result showed that the divergences of *rep* and *oriV* region were considerable. In their study, 30 IncP-9 plasmids had up to 35% divergence by phylogenetic analysis. However, cultivation-independent methods to study the diversity and abundance of IncP-9 in the more complex environmental samples such as soils have only rarely been performed [1].

In this study, we developed a new primer set based on 28 *oriV* and *rep* sequences of IncP-9 plasmids. This primer pair was checked *in silico* and experimentally with plasmid DNA belonging to IncP-9 and other groups. Finally the new primers were applied to study the two soil samples with/without phenanthrene contamination.

Material, methods, result and discussion

To detect a wide spectrum of targeted genes, degenerated primers are frequently used. But too high degeneracy could lead to low efficiency in amplification or distortion of community structure. How to reduce primers degeneracy without narrowing their detection spectrum? We proposed a strategy for primer design which follows two rules: (1) degeneracies of each primer are allowed for the last 6 bp at 3' end of the primer to perfectly match the target sequences; (2) mismatches at 5' end of the primers were introduced in an attempt to reach an equal amount of mismatch for each targeted sequences type. Alignments of the *oriV-rep* region of twenty-eight sequences of *oriV* (EU499619-EU499641, AF078924, AB237655, AJ344068, AB257759 and AF491307) and *rep* (EU499644-EU499666, AF078924, AB237655, AJ344068, AB257760 and AF491307) were used. Multiple alignments were performed with MEGA. The selected primer system consists of 21-mer degenerated forward primer (5-GAG GGT TTG GAG ATC ATW AGA-3) and reverse primer (5-GGT CTG TAT CCA GTT RTG CTT-3). *In silico* analysis showed no mismatch for at least 12 bp at the 3'end of each primer and 1-4 mismatches for each sequence type at the 5'end (Fig. S1). The expected amplicon size is 610-637 bp. Optimization of the primers was done with one genomic DNA with plasmid pNF142 and one TC DNA extracted from artificial contaminated soil samples at two annealing temperatures (52-53°C). The primers were further tested with plasmid DNA summarized in Table 1. None of the plasmids belonging to other incompatibility groups was amplified while the reference plasmids were amplified. PCR amplification was performed in 25 µl reaction mixture containing 1X Stoffel buffer, 0.2 mM deoxynucleoside, 2.5 mM $MgCl_2$, 1 µg/µl bovine serum albumin, 0.2 µM of each primer, and 2.5 U Amplitaq DNA polymerase. Denaturation was carried out at 94°C for 5 min, followed by 35 cycles of 1 min at 94°C, 1 min at 53°C and 1.5 min at 72°C. The PCR finally extended at 72°C for 10 min.

The setup of the experiment was done as follows: briefly soils were contaminated with phenanthrene to reach a final concentration of 2 mg g^{-1} soil and incubated at room temperature (23 °C). Samples were taken at days 0, 21 and 63 from the four soil microcosms per treatment and kept at -20 °C before DNA extraction with Bio-101 DNA spin kit for soil (Qbiogene, Heidelberg, Germany). The PCR system was applied

to study the response of IncP-9 harboring bacteria to phenanthrene contamination in soil samples collected at day 63. Details on the experiment, soil parameters, phenanthrene analysis, quantitative real-time PCR (qPCR) for 16S rRNA gene, studies on abundance and diversity of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes and soils' sorption ability on phenanthrene were described previously [14]. The response of microbial communities was studied and described in chapter 3. Amplicons with expected size were acquired from all samples spiked with/without phenanthrene. The bands for Luvisol samples were weak and accompanied by unspecific amplicons possibly due to the low abundance of target sequences in TC-DNA. The amplicons were further checked by Southern blot analysis using a mixed probe generated from typical IncP-9 plasmids (pM3, pBS265, pMG18, pNL60, pNL15, pSVS15, pNF142, pSN11, P80 and pWWO) representing different IncP-9 subgroups. Interestingly, one replicate each of phenanthrene-treated Cambisol, phenanthrene-contaminated Luvisol and the control of Luvisol did not hybridize with the mixed probes. This result suggests not only new *oriV-rep* regions may be present in the soil samples, but also a highly heterogeneous distribution of IncP-9 plasmids. However, no increased abundance of IncP-9 harboring bacteria in response to phenanthrene spiking was detected for both soils (Fig. 1a—d). This result nicely confirmed findings from a previously published study on PAH-*RHDα* genes using the same DNA samples. In this work *phnAc* genes were shown to be dominant in the contaminated Luvisol while a high diversity of *PAH-RHDα* genes of Gram-positive and Gram-negative bacteria in the phenanthrene-spiked Cambisol. But none of the detected *PAH-RHDα* genes could be affiliated to *nahAc* genes which are typically harbored on IncP-9 plasmid. Because the identification of microbial responders to phenanthrene contamination is of priority in this project, the amplicons of *oriV-rep* regions were not subjected for further analysis. Therefore, the successful application and further development of the primer system was achieved in the analysis of total community DNA from biofilter samples. Cloning and sequencing of IncP-9 *oriV-rep* PCR amplicons confirmed the specificity of the primers and a large diversity of IncP-9 plasmids in biofilters. In conclusion, so far the PCR system is suitable for studying the diversity and abundance in environmental samples for IncP-9.

ACKNOWLEDGMENTS

This work was supported by DFG SPP1315 (SM59/8–1).

Refenece

1. Sevastsyanovich YR, Krasowiak R, Bingle LE, Haines AS, Sokolov SL, et al. (2008) Diversity of IncP-9 plasmids of Pseudomonas. Microbiology 154: 2929-2941.

2. Lambertsen LM, Molin S, Kroer N, Thomas CM (2004) Transcriptional regulation of pWW0 transfer genes in *Pseudomonas putida* KT2440. Plasmid 52: 169-181.

3. Greated A, Titok M, Krasowiak R, Fairclough RJ, Thomas CM (2000) The replication and stable-inheritance functions of IncP-9 plasmid pM3. Microbiology 146 (Pt 9): 2249-2258.

4. Heinaru E, Vedler E, Jutkina J, Aava M, Heinaru A (2009) Conjugal transfer and mobilization capacity of the completely sequenced naphthalene plasmid pNAH20 from multiplasmid strain *Pseudomonas fluorescens* PC20. FEMS Microbiol Ecol 70: 563-574.

5. Sota M, Yano H, Ono A, Miyazaki R, Ishii H, et al. (2006) Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase. J Bacteriol 188: 4057-4067.

6. Greated A, Lambertsen L, Williams PA, Thomas CM (2002) Complete sequence of the IncP-9 TOL plasmid pWW0 from *Pseudomonas putida*. Environ Microbiol 4: 856-871.

7. Schneiker S, Keller M, Droge M, Lanka E, Pühler A, et al. (2001) The genetic organization and evolution of the broad host range mercury resistance plasmid pSB102 isolated from a microbial population residing in the rhizosphere of alfalfa. Nucleic Acids Res 29: 5169-5181.

8. Izmalkova TY, Mavrodi DV, Sokolov SL, Kosheleva IA, Smalla K, et al. (2006) Molecular classification of IncP-9 naphthalene degradation plasmids. Plasmid 56: 1-10.

9. Krasowiak R, Smalla K, Sokolov S, Kosheleva I, Sevastyanovich Y, et al. (2002) PCR primers for detection and characterisation of IncP-9 plasmids. FEMS Microbiol Ecol 42: 217-225.

10. Gomes NC, Flocco CG, Costa R, Junca H, Vilchez R, et al. (2010) Mangrove microniches determine the structural and functional diversity of enriched petroleum hydrocarbon-degrading consortia. FEMS Microbiol Ecol 74: 276-290.
11. Greated A, Thomas CM (1999) A pair of PCR primers for Incp-9 plasmids. Microbiology 145 (Pt 11): 3003-3004.
12. Krasowiak R, Sevastsyanovich Y, Konieczny I, Bingle LE, Thomas CM (2006) IncP-9 replication initiator protein binds to multiple DNA sequences in oriV and recruits host DnaA protein. Plasmid 56: 187-201.
13. Sevastsyanovich YR, Titok MA, Krasowiak R, Bingle LE, Thomas CM (2005) Ability of IncP-9 plasmid pM3 to replicate in *Escherichia coli* is dependent on both rep and par functions. Mol Microbiol 57: 819-833.
14. Ding GC, Heuer H, Zuhlke S, Spiteller M, Pronk GJ, et al. (2010) Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes by using a novel PCR detection system. Appl Environ Microbiol 76: 4765-4771.

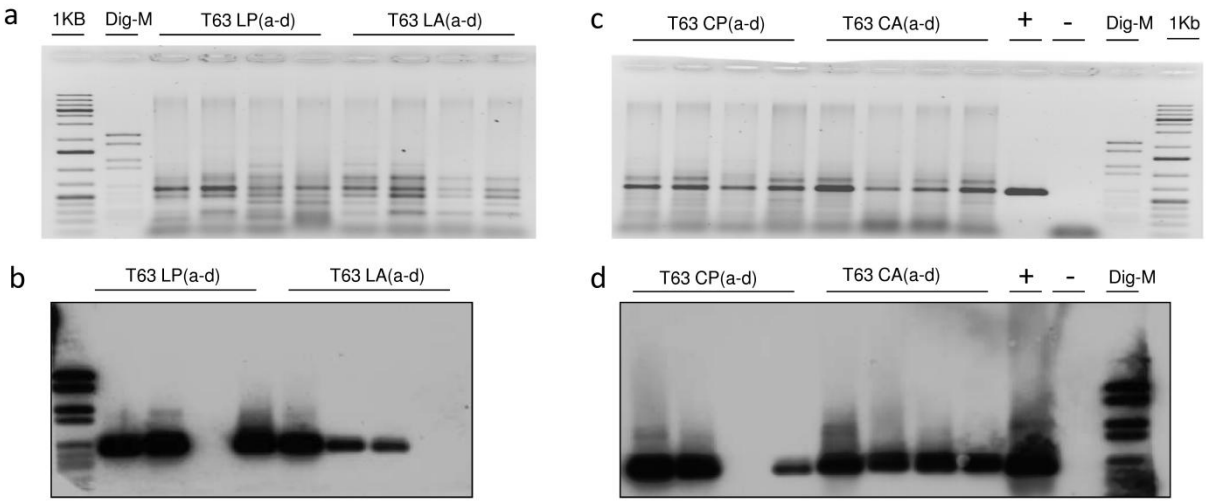

Fig 1: PCR (a, c) and southern blot (b, d) analyses of oriV-rep regions of incP-9 plasmids for soil spiked with phenanthrene (LP, CP) or not (LA, CA) at days 63.

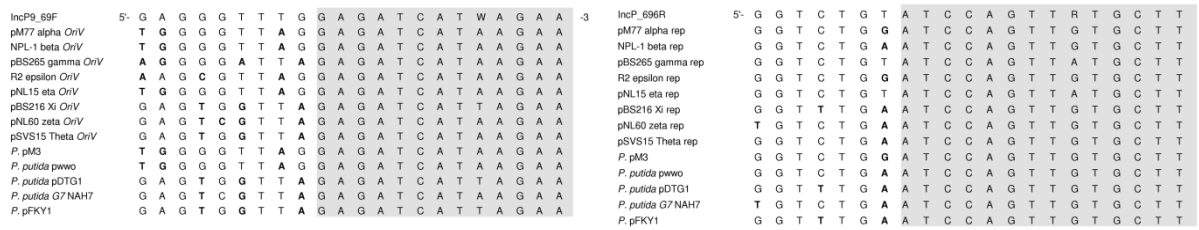

Fig. S1: Alignment of primers and typical oriV regions and *rep* genes of incP9 plasmids

Chapter 6: Similar bacterial populations increased in two different phenanthrene spiked soils from distant geographic sites as revealed by DGGE and pyrosequencing of 16S rRNA gene fragments

Guo-Chun Ding, Holger Heuer and Kornelia Smalla*

Julius Kühn-Institut (JKI) – Federal Research Centre for Cultivated Plants, Messeweg 11-12, 38104 Braunschweig, Germany

Running title: Soil microbial community response to phenanthrene

Key words: soil, phenanthrene, total community DNA, 16S rRNA gene, DGGE, pyrosequencing, *Sphingomonadales*, *Burkholderiales*

Corresponding author:

Prof. Dr. Kornelia Smalla, JKI, Messeweg 11-12, D-38104 Braunschweig, Germany

kornelia.smalla@jki.bund.de; Phone: +49-531-2993814; FAX: +49-531-2993006

Abstract

To explore the effect of polycyclic aromatic hydrocarbons on indigenous microbial communities, phenanthrene was added as a model pollutant in a microcosm experiment with two different soils. Total community (TC) DNA was extracted from samples collected after 21 and 63 days of phenanthrene addition. The microbial communities were studied by denaturing gradient gel electrophoresis (DGGE) fingerprinting analysis of 16S rRNA gene or ITS fragments amplified from TC DNA and 16S rRNA amplicon pyrosequencing. By DGGE, soil type, taxonomic group and exposure time dependent responses to phenanthrene spiking were observed for bacterial communities. *Polaromonas* sp. was identified as most dominant betaproteobacterial responders in both soils by sequencing DGGE bands with increased intensity. In contrast to bacteria, no changes in the fungal communities were detected in the phenanthrene spiked soils. Pyrosequencing analysis of the 16S rRNA V3-V4 region amplified from day 63 DNA revealed distinct soil type and treatment-dependent bacterial community structure and several genera such as *Sphingomonas, Phenylobacterium, Polaromonas, Peredibacter, Mycobacterium* and *Mucilaginibacter* that were significantly enriched in both polluted soils. Strikingly, 30 out of 88 operational taxonomic units (>97% similarity) were enriched in both contaminated soils originating from distant sites.

INTRODUCTION

The fate of polycyclic aromatic hydrocarbons (PAHs) in soil is of great interest due to their potential for bioaccumulation, persistence, transport and toxicity (19, 25-26). Microbes play an important role in cleaning up PAH-polluted sites. Cultivation-based approaches were often used to isolate PAH degraders from contaminated environments or after PAH enrichments. Based on 16S rRNA gene sequencing these PAH degraders were frequently reported to be affiliated to several genera such *as Sphingomonas, Polaromonas, Burkholderia, Pseudomonas, Mycobacterium, Nocardia, Rhodococcus* (1, 5, 8, 14, 28, 33, 48). It is reasonable to assume that bacteria belonging to these taxa are PAH degraders when they enriched after exposure to PAH pollution. But soil is a complex environment where non PAH degrader could also contribute to PAH degradations. For example, bacteria which produce bio-surfactants could enhance the bioavailability of hydrophobic compounds such as PAHs, and microbes with the ability to degrade low molecular weight aromatic hydrocarbons could reduce the toxicity of intermediates of PAH degradation (35). Fungi were also reported to be involved in PAH remediation by secreting peroxidases or transporting bacterial degraders to the contaminated sites (7, 29, 36). Whether this putative involvement could change the relative abundance of specific fungi is still unknown.

However, it might be too simple to assume that bacteria or fungi enriched in response to PAH pollution are exclusively PAH degraders when studying complex environments such as soils. To achieve a better understanding of the cooperative mineralization network for PAHs in soils, it is important to overcome the biases of cultivation-based enrichments. Molecular analysis of total community DNA might reveal broader insights into the complex and cooperative response of soil microbial communities to PAH pollution and allow identifying responders to the PAH pollution. Under aerobic conditions, microbial mineralization of PAHs is initiated by dioxygenases which incorporate molecular oxygen into the aromatic ring. Genes encoding these enzymes were previously used as genetic markers to study the responses of the bacteria which initiate PAH mineralization (3, 4, 10, 13, 32, 45). PCR-based methods employing primers targeting ring hydroxylating dioxygenase (RHD) genes provided insights into their diversity, distribution and abundance in different environments. Several studies aimed to correlate the abundance of ring

hydroxylating dioxygenase genes with PAH pollution. However, PAH pollution might affect not only bacterial populations that carry *PAH-RHD* genes but a much wider network.

To study the changes of the soil microbial community structure in response to PAHs, microcosm experiments were performed with typical Luvisol (silt loam) and Eutric Cambisol (clay loam) using phenanthrene as model pollutant. Recently, Ding et al. (9) reported on the soil type dependent abundance and diversity of *RHDα* genes in phenanthrene spiked Luvisol and Cambisol (9). The authors employed a cultivation-independent approach using novel primers that target *RHDα* genes of both Gram-positive and Gram-negative bacteria. It was hypothesized that a soil type dependent microbial community composition as well as the different soil characteristics (mineral composition, organic matter, pore size) might have contributed to the soil type dependent response of microbial communities to phenanthrene spiking. The goal of this study was to analyze TC DNA from the same experiment to elucidate shifts in the relative abundance of bacterial and fungal populations in response to phenanthrene spiking by using a cultivation-independent 16S rRNA gene or ITS fragment based approach. In order not only to analyze the dominant bacterial community members a set of different group-specific primers was employed. DGGE fingerprints of *Bacteria, Proteobacteria*-specific groups, *Actinobacteria* or *Fungi* were used to monitor overall shifts of the microbial community in two different soils after phenanthrene spiking. This approach provided the chance to improve our understanding of the complex network of bacteria and fungi responding to PAH pollution in soils. In addition, pyrosequencing of 16S rRNA gene amplicons was used to study the composition of the bacterial community in both soils and to identify responding taxa to the soil spiking with phenanthrene.

MATERIALS AND METHODS

Experimental design. Details of the experimental setup, soil parameters, phenanthrene analysis, quantitative real-time PCR (qPCR) for 16S rRNA genes, the abundance and diversity of *PAH- RHD* genes and soils' sorption ability on phenanthrene were described previously (9). Briefly, Cambisol and Luvisol soil samples taken from the long-term observation sites in Ultuna and Scheyern, respectively, were contaminated with phenanthrene to reach a final concentration of

2 mg g^{-1} soil and incubated at room temperature (23 °C). Samples were taken at days 0, 21 and 63 from the four replicate soil microcosms per treatment and kept at -20 °C before DNA extraction with Bio-101 DNA spin kit for soil (Qbiogene, Heidelberg, Germany). The same DNA samples were used as basis for both studies.

DGGE analysis of 16S rRNA gene fragments and ITS1. Bacterial 16S rRNA gene fragments from the soil samples were amplified with the primers F984-GC and R1378 as described by Heuer et al. (2). A nested PCR approach was applied for amplification of 16S rRNA genes of Actinobacteria and Proteobacteria specific groups as previously described (6, 16, 21). For all nested PCRs, the first amplification was of 30 thermal cycles and then the amplicons were checked on agarose gel to make proper dilutions of templates for the second amplification with 25 thermal cycles. DGGE of the 16S rRNA gene amplicons was performed according to Gomes et al. (14). Amplification of fungal ITS regions and the following DGGE were done according to Weinert et al. (44). The DGGE gels were silver stained according to Heuer et al. (22). Software package GelCompare 4.5 was used for cluster analysis of DGGE profiles. Dendrograms were constructed based on pair wise Pearson correlation indices within a gel using unweighted pair group method using arithmetic averages (UPGMA). The pair wise Pearson correlation index was also subjected to permutation tests using the PERMTEST software developed by Kropf et al. (30).

Cloning and sequencing of the selected bands from DGGE gels were performed according to Gomes et al. (15). The partial 16S rRNA gene sequences excluding the primers were subjected to CHIMERA CHECK program from RDP and then classified by the Naïve Bayesian rRNA Classifier of the Ribosomal Database Project (http://rdp.cme.msu.edu/) and BLASTN in the GenBank database (http://blast.ncbi.nlm.nih.gov)

Pyrosequencing and sequence analyses. Bacterial communities of samples collected at day 63 were further studied by pyrosequencing with three replicates per treatment. PCR and sequencing were done at Biotechnology Innovation Center (BIOCANT) in Portugal. Briefly, hypervariable V3-V4 regions of 16S rRNA genes were amplified with bacterial primers, 338F (43) and 802R (RDP's Pyrosequencing

Pipeline: http://pyro.cme.msu.edu/pyro/help.jsp) which was fused to the 454 A and B adaptors, respectively. Standard PCR reaction conditions were employed for reactions with Fast Start polymerase (Roche, Pensberg, Germany), 3 mM $MgCl_2$, 6% DMSO, 200 nM each primer and 200 mM dNTP. The PCR conditions were 94 ºC for 3 min, followed by 35 cycles of 94 ºC for 30 s, 44 ºC for 45 s and 72 ºC for 60 s and a final elongation step at 72 ºC for 2 min. Sequencing was performed on a 454 Genome Sequencer FLX platform according to standard 454 protocols (Roche – 454 Life Sciences, Branford, CT, USA).

To remove the noise data, all sequences were analyzed with standalone BLASTN tools from NCBI against a local 16S rRNA gene database. Frequently, non high-scoring segment pairs (HSP) were found at the tail region. Therefore, for each sequence the unpaired region was truncated and the remaining part was checked whether its length is above 200 bp. Software package Mothur (v1.14.0) was used for multiple alignment, OTU assignment and rarefaction analysis (38-39). Multiple alignments were done based on the Silva bacterial database (37) and OTUs were assigned at 3% DNA distance. Sequences were classified by Naïve Bayesian classifier in Ribosomal database project at >80% confidence (42). Aligned sequences, their corresponding taxonomy, as well as OTU assignment were stored in a local MYSQL database. Self-written Perl script including the package database interface (DBI) was used to retrieve sequences from the local database and construct an OTU level report with each row representing one OTU containing taxonomic position (domain, phylum, class, order, family, genus) and number of sequences for each sample. Data summary at different taxonomic levels, cluster analysis based on pair wise Pearson Correlation distance and plot of rarefaction curves were performed with software R (2.12) (http://www.r-project.org/). To identify the taxa which were significantly enriched in the polluted soils or significantly differed between soil types, multiple Tukey's honest significance tests under a generalized linear model via a logistic function for binomial data were done with the R add-on package multcomp (24). To avoid the errors for the statistical model, the number of detected sequences for one replicate was shifted to 1 if certain taxa were not detected for the treatments. Representative sequences for those OTUs with significantly higher relative abundance in phenanthrene spiked soil were retrieved from the local database. Based on BLASTN reports, their corresponding closest related 16S rRNA gene

sequences in RDP databases were also retrieved. After multiple alignments with Mothur, corresponding regions (V3-V4) were selected for phylogenetic analysis. A neighbor joining phylogenetic tree was constructed based on pair wise Jukes & Cantor DNA distances and the tree was tested by 500 times bootstrap analysis using software package SeaView4 (17). Phylogenetic trees were edited with software Archaeopteryx (20).

RESULTS

Responses of soil bacterial communities to phenanthrene contamination depended on soil type, exposure time and taxonomic group as revealed by PCR-DGGE. Community fingerprints of Bacteria, Pseudomonas, Betaproteobacteria, Actinobacteria, Alphaproteobacteria and Fungi were generated and the profiles were clustered by UPGMA based on Pearson correlation indices (Fig. S1-S6). Shifts in the bacterial community patterns were more pronounced in Luvisol than in Cambisol at both sampling times (Fig. S1a). Distinct treatment-dependent clusters were observed at both time points (Fig. S1b). The similarities between the polluted soils and the corresponding controls decreased with time. Several bands with strongly increased intensity were detected in the fingerprints of the phenanthrene-treated soils. Permutation tests confirmed significant effects of phenanthrene treatments for both soils (Table 1). Despite the relatively high variability of the Pseudomonas-specific gacA fingerprints some differentiating bands were already detected in the patterns of phenanthrene-polluted Luvisol and Cambisol at day 21 (Fig. S2a, b). At day 63, no differentiating bands were identified in response to the phenanthrene contamination in the Pseudomonas fingerprints (Fig. S2c), and although the fingerprints of both soils formed distinct clusters no treatment-dependent clusters were observed (Fig. S2d). In contrast to *Pseudomonas* communities, shifts of *Actinobacteria* in response to phenanthrene pollution were observed only at day 63 but not at day 21 (Fig. S3a-d). Based on the cluster analysis the shifts in phenanthrene-polluted soils were stronger for Luvisol than for Cambisol (Fig. S3d). Interestingly, the Betaproteobacteria patterns of both phenanthrene-polluted soils were highly similar at both time points, and the populations (accession number: JF810414-JF810419) behind the bands with strongly increased intensity (most dominant responder) from both soils shared the highest sequence similarity with *Polaromonas* sp. (AM 492164) (Fig. S4a-c). Permutation tests confirmed the significant effect of phenanthrene contamination on

the betaproteobacterial community at both sampling times and for both soils (Table 1). The alphaproteobacterial community patterns for soil samples taken at day 21 displayed a high variability among replicates and although some differentiating bands were identified when comparing polluted and unpolluted Luvisol, the treatment effects were not significant as suggested by Permutation tests for both soils (Table 1). Neither clear soil type nor treatment-dependent clusters were observed at day 21. At day 63 soil type and treatment-dependent clusters and significant differences were observed between polluted and non-polluted treatments for both soils (Fig. S5a,b).
In contrast to bacterial communities, fungal communities did not show clear responses to phenanthrene contamination at both sampling times though distinct clusters for both soil types were formed (Fig. S6a-d). Although no treatment-dependent bands were identified, the permutation test revealed significant treatment effects for the Cambisol.

Distinct soil type- and treatment dependent bacterial community composition revealed by 16S rRNA amplicon sequencing. Pyrosequencing of 16S rRNA genes (V3-V4) amplified from total community DNA of soil samples taken 63 days after phenanthrene spiking was performed for three replicates per treatment. Sequences were classified by the Naïve Bayesian classifier of the Ribosomal database project. Altogether 34,054 sequences for all 12 samples were examined, of which 31,201 sequences could be classified to 21 phyla. The remaining 2,853 sequences were only classified as *Bacteria*. Most sequences could be classified at the higher taxonomic levels (phyla, class) while still more than 40% sequences could not be reliably classified at genus level indicating novel not yet described diversity (Table 2). In comparison to the control soil, lower numbers of sequences were obtained for both polluted soils. The most abundant taxa in both soils belonged to the *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Firmicutes*, and *Gemmatimonadetes* (Fig. 1). In both polluted soils the relative abundance of sequences belonging to the *Proteobacteria* was much higher compared to the control (Fig. 1). To compare the diversity of detected sequences from different treatments, rarefaction analyses were performed based on OTUs defined at four similarities (97%, 90%, 80% and 70%). Higher detected diversities were found for non-polluted soils compared to the corresponding polluted soils at all similarity levels (Fig. 2). The highest diversity was detected for the control Luvisol. The rarefaction curves based on OTUs assigned at >97% similarity

did not reach a plateau while the rarefaction curves of OTUs defined at >90%, 80% and 70% did. This result suggested that more sequences are still needed to represent the diversity of soil samples at the species levels while the dominant phyla, classes, families and genera have already been detected. Bacterial community structure analyses based on the OTU reports (>97% and >80%) confirmed the result of the DGGE analysis (Fig. 3a,b). Separated clusters were observed for all treatments based on OTUs defined at 97% similarity. Interestingly, the polluted soils clustered together when using OTUs defined at >80% similarity indicating that at higher taxonomic level similar OTUs responded to the phenanthrene pollution.

Common taxa enriched in both phenanthrene contaminated soils as revealed by amplicon pyrosequencing analysis. To identify those taxa which were enriched in the polluted soils, multiple Tukey's tests were applied (unadjusted $p<0.05$). Genera belonging to the *Alphaproteobacteria* (*Novosphingobium, Sphingomonas, Phenylobacterium*, *Afipia*), the *Betaproteobacteria* (*Polaromonas, Burkholderia, Rhodanobacter*), *Deltaproteobacteria* (*Peredibacter*), *Actinobacteridae (Mycobacterium)* and *Bacteroidetes* (*Mucilaginibacter*) were found enriched in both polluted soils (Table 3). In particular the genera *Sphingomonas* and *Polaromonas* were found to be significantly enriched in both phenanthrene-polluted soils. Although *Sphingomonas* was the dominant genus in the unpolluted Cambisol and Luvisol with 3% and 4%, respectively, its relative abundance strikingly increased to 23% and 18% in phenanthrene spiked soils (Table 3). In contrast, *Polaromonas* was rarely obtained from both unpolluted soils but the proportion of sequences assigned to this genus drastically increased in the phenanthrene-polluted soils to 4% and 7% in Cambisol and Luvisol, respectively. Soil type dependent responses were also observed. For example, *Sphingosinicella, Herbaspirillum*, *Dokdonella*, *Dyella*, *Aquicella*, *Legionella* and *Dactylosporangium* were enriched in the polluted Luvisol while *Bosea*, *Duganella*, *Pseudoxanthomonas* and *Pseudomonas* were enhanced in the polluted Cambisol. To pinpoint those species which significantly increased in phenanthrene-polluted soils, sequences were assigned to OTUs at 3% DNA distance. The majority of OTUs enriched in both soils belonged to the *Proteobacteria*. A total of 66 OTUs with significantly increased abundance in the polluted Luvisol soil were affiliated to *Proteobacteria*. The number of enriched OTUs in the polluted Luvisol (69 OTUs) was higher than that in the polluted Cambisol (49 OTUs). Although fewer OTUs with

significantly increased abundance were observed in polluted Cambisol, these OTUs belonged to four different phyla. Most strikingly, 30 out of 88 OTUs were enriched in both polluted soils collected from distant sites (Table 4). The majority of these commonly enriched OTUs belonged to *Sphingomonadaceae* (22 OTUs) and *Polaromonas* (4 OTUs). Soil type specific OTUs responding to phenanthrene spiking were also observed for both soils. OTUs which were only enriched in the polluted Luvisol could be affiliated exclusively with proteobacterial classes *Alphaproteobacteria* (44 OTUs), *Betaproteobacteria* (11 OTUs), *Gammaproteobacteria* (7 OTUs) and *Deltaproteobacteria* (4 OTUs) with the exception of two OTUs to *Actinobacteria* and one OTU that could not be affiliated to any phylum. OTUs which were only enhanced in the polluted Cambisol (19 OTUs) were affiliated with *Proteobacteria* (44 OTUs), *Actinobacteria* (3 OTUs), *Acidobacteria* (1 OTU) and *Bacteroidetes* (1 OTU) (Fig. 4 and Table 4).

DISCUSSION

In order to achieve a better understanding of the complex microbial community network response in two different phenanthrene spiked soils, changes in the bacterial community composition were investigated by DGGE and amplicon pyrosequencing analyses of 16S rRNA gene fragments amplified from total community DNA. DGGE analysis revealed sampling time, soil type and treatment dependent shifts in the different bacterial taxa analyzed. Over time the bacterial community composition of polluted soils and control soils became more dissimilar for all taxa analyzed except for *Pseudomonas* while the similarity of the polluted soils increased. These data indicate similar changes in the relative abundance of bacteria in both soils which most likely resulted from growth of bacteria utilizing phenanthrene or its metabolites. Interestingly, for some taxa, e.g. *Actinobacteria*, shifts in the relative abundance were only observed at Day 63. This might be due to the relative low abundance of PAH mineralizing populations such as *Mycobacterium* in both soils and the slower growth rates of populations belonging to this genus. Bands with enhanced intensities were observed for all bacterial community profiles of all taxa analyzed at one or both sampling times for both polluted soils.

The shifts in the DGGE patterns observed for Luvisol after phenanthrene pollution were more pronounced compared to Cambisol. Ding et al. (9) also reported on a more increase in the abundance of bacterial populations carrying PAH-RHD*α* genes in phenanthrene spiked Luvisol. In that study, 21 days after phenanthrene spiking PAH-RHD gene PCR products were only obtained from total community DNA of Luvisol but not from the identically treated Cambisol. At 63 days after spiking, PCR products were detected in total community DNA from both polluted soils but not in the control soils. Quantitative real-time PCR revealed that the relative abundance of PAH-RHD genes was more than one order of magnitude higher for Luvisol compared to Cambisol, and also indicated a significant increase of 16S rRNA gene copy numbers in polluted Luvisol compared to the control soil (9). An increase of the 16S rRNA gene copy numbers, although not significant, was also observed in the phenanthrene spiked Cambisol (9). Differences in the soil texture might have contributed to the stronger response of the microbial community to phenanthrene spiking in the Luvisol compared to Cambisol as evidenced by significantly increased 16S rRNA gene copy numbers and more pronounced DGGE shifts in Luvisol (9). The clay content in the Luvisol (14%) was much lower than that in the Cambisol (37%). While Müller and Höper (35) showed a negative relationship between soil clay content and microbial respiration activity, Uettebroek et al (41) demonstrated an increased phenanthrene mineralization in the clay fraction that correlated with a higher relative abundance of *Mycobacteria* in this fraction. The differences in the initial microbial community composition of the Luvisol and the Cambisol soil might be another reason. The analysis of 16S rRNA gene fragments by DGGE or pyrosequencing revealed a distinct bacterial community composition of both soils. Despite a large number of OTUs which were enriched from both polluted soils, some enriched OTUs were soil type specific, e.g. two OTUs which were only detected in Luvisol and displayed a high sequence similarity to *Byssovorax cruenta* (AJ833647) and *Nannocystis sp* (AY996823) (Fig. 2).

Although several studies suggested a role of fungi in PAHs metabolization by direct degradation, by transport of bacterial degraders to the source of PAHs, or by enhancing phytoremediation (11, 23, 29, 34, 49), no shifts in the community profiles for fungi were detected in both polluted soils suggesting that the relative abundance of dominant fungal populations did not change. It can, however, not be excluded that

mixing the phenanthrene into soil might have impaired existing fungal networks. Nevertheless, chemical analysis showed that the concentration of phenanthrene in both soils at Day 63 were very low (9) and therefore fungal networks might have still played an important role in transporting both phenanthrene and the bacterial degraders.

Strikingly responders most likely belonged to the rare biosphere became detectable in both soils after phenanthrene spiking. The analysis of 16S rRNA gene amplicon pyrosequence data revealed that several genera such as *Sphingomonas, Phenylobacterium, Polaromonas, Peredibacter, Mycobacterium* and *Mucilaginibacter* were of significantly higher relative abundance in both polluted soils compared to the corresponding control soils. Rarefaction analyses suggested a decreased richness for both phenanthrene-polluted soils. But the lower diversity observed for the polluted soils resulted most likely from the increased abundance of few OTUs in response to phenanthrene spiking. These enriched bacteria were affiliated with different phyla such as *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes*. However, several genera enriched in polluted soils were not known as PAH degraders, e.g. *Sphingosinicella*, *Afipia*, *Duganella*, *Herbaspirillum*, *Rhodanobacter*, *Dokdonella*, *Dyella*, *Aquicella*, *Peredibacter*, *Legionella*, *Nannocystis*, *Byssovorax*, *Dactylosporangium*, and *Mucilaginibacter*. Bacteria belonging to the genera *Sphingomonas* and *Polaromonas* were dominant responders in both polluted soils. Remarkably, the 16S rRNA gene sequence behind the betaproteobacterial DGGE band with increased abundance in the polluted Luvisol and Cambisol also shared high sequence similarity with *Polaromonas* (>95%). Isolates belonging to the genera *Sphingomonas* and *Polaromonas* were reported in several studies as PAH degraders (18, 31, 46). Alphaproteobacterial and betaproteobacterial 16S rRNA gene sequences were also found as dominant responding populations in PAH polluted sandy peat soils based on 16S rRNA gene sequencing a total of 39 clones (47). Gomes et al. (2007) used DGGE to follow shifts in the relative abundance of soil bacteria in response to naphthalene (13). In this study the populations behind bands with increased abundance (responders to the pollutant) were identified as *Burkholderia* sp. and *Rhodococcus*.

Alphaproteobacteria (Sphingomonadaceae) *RHDα*-like genes were detected from both polluted soils but not from the corresponding control soils. Quantitative real-time

PCR revealed that the copy number of *RHDα*-like genes related to 16S rRNA gene copy number was estimated to be less than 1% in TC-DNA of polluted soils (9). However, pyrosequence analyses suggested that 23% and 18% of the sequences detected for the polluted Cambisol and Luvisol were affiliated to *Sphingomonas,* respectively. It is also important to note that a high number of sequences obtained from control soils were also affiliated to *Sphingomonas.* To exclude that multiple copies of 16S rRNA genes per *Sphingomonas* genome might lead to an overestimation of the relative abundance of *Sphingomonas*, an *in silico* analysis was performed to find out the number of copies of 16S rRNA genes based on the complete genomes of *Sphingomonadaceae* (7) in GenBank (19/04/2011). On average, two copies of 16S rRNA genes per genome were found. Therefore, pyrosequence analyses showed that bacteria belonging to *Sphingomonas* are dominant in both polluted and unpolluted soils. However, only a small proportion of *Sphingomonas* populations seem to carry *RHD*□-like genes that are essential for the first step of the aerobic phenanthrene degradation. The majority of *Sphingomonas* populations might have profited from metabolites of phenanthrene degradation. In addition, the high initial relative abundance of *Sphingomonas* probably also contributed to the dominance in the phenanthrene polluted soils.

The model soils which were classified as Cambisol (Ultuna, Sweden) and Luvisol (Scheyern, Germany) originated from distant geographic sites in Northern Europe. Distinct bacterial community structure and soil type dependent abundance and diversity of PAH-RHD□ genes in response to phenanthrene spiking were detected. Surprisingly, twenty-three common OTUs (>97% similarity) were enriched in both polluted soils collected from distant sites. These findings indicate that certain bacterial functional groups are widely distributed. In the study of Fulthorpe et al. (2008) few common species were also found from different soils sampled from distant sites (12).

In summary, this study indicated a complex mineralization network for phenanthrene in soil. Both DGGE and pyrosequence analysis of 16S rRNA gene fragments amplified from total community DNA proved that phenanthrene spiking induced fundamental changes of the bacterial community structure of both soils. Numerous populations were enriched in both polluted soils that belonged to genera not previously described as PAH degraders. Their role in the complex network remains to be elucidated. The 16S rRNA (this study) and the PAH-RHA□ based

analysis of bacterial communities in phenanthrene spiked soils clearly demonstrated the importance of the rare biosphere as PAH-RHAα genes and many of the enriched populations were only detectable in the in phenanthrene spiked soils.

ACKNOWLEDGMENTS

This work was supported by DFG SPP1315 (SM59/8–1) and the BMBF project MÄQNU.

REFERENCES

1. **Andreoni, V., S. Bernasconi, M. Colombo, J. B. van Beilen, and L. Cavalca.** 2000. Detection of genes for alkane and naphthalene catabolism in *Rhodococcus* sp. Strain 1BN. Environ. Microbiol. **2:**572-577.
2. **Brim, H., H. Heuer, E. Krögerrecklenfort, M. Mergeay, and K. Smalla.** 1999. Characterization of the bacterial community of a zinc-polluted soil. Can. J. Microbiol. **45:**326-338.
3. **Cebron, A., B. Louvel, P. Faure, C. France-Lanord, Y. Chen, J. C. Murrell, and C. Leyval.** 2011. Root exudates modify bacterial diversity of phenanthrene degraders in PAH-polluted soil but not phenanthrene degradation rates. Environ. Microbiol. **13:**722-736.
4. **Cebron, A., M. P. Norini, T. Beguiristain, and C. Leyval.** 2008. Real-Time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHDalpha) genes from Gram positive and Gram negative bacteria in soil and sediment samples. J. Microbiol. Methods **73:**148-159.
5. **Chain, P. S., V. J. Denef, K. T. Konstantinidis, L. M. Vergez, L. Agullo, V. L. Reyes, L. Hauser, M. Cordova, L. Gomez, M. Gonzalez, M. Land, V. Lao, F. Larimer, J. J. LiPuma, E. Mahenthiralingam, S. A. Malfatti, C. J. Marx, J. J. Parnell, A. Ramette, P. Richardson, M. Seeger, D. Smith, T. Spilker, W. J. Sul, T. V. Tsoi, L. E. Ulrich, I. B. Zhulin, and J. M. Tiedje.** 2006. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. PNAS **103:**15280-15287.

6. **Costa, R., N. C. Gomes, E. Krögerrecklenfort, K. Opelt, G. Berg, and K. Smalla.** 2007. *Pseudomonas* community structure and antagonistic potential in the rhizosphere: insights gained by combining phylogenetic and functional gene-based analyses. Environ. Microbiol. **9:**2260-2273.
7. **Dai, C. C., L. S. Tian, Y. T. Zhao, Y. Chen, and H. Xie.** 2010. Degradation of phenanthrene by the endophytic fungus Ceratobasidum stevensii found in Bischofia polycarpa. Biodegradation **21:**245-255.
8. **Dean-Ross, D., and C. E. Cerniglia.** 1996. Degradation of pyrene by *Mycobacterium flavescens* Appl. Microbiol. Biotechnol. **46:**307-312.
9. **Ding, G. C., H. Heuer, S. Zühlke, M. Spiteller, G. J. Pronk, K. Heister, I. Kögel-Knabner, and K. Smalla.** 2010. Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes by using a novel PCR detection system. Appl. Environ. Microbiol. **76:**4765-4771.
10. **Flocco, C. G., N. C. Gomes, C. W. Mac, and K. Smalla.** 2009. Occurrence and diversity of naphthalene dioxygenase genes in soil microbial communities from the Maritime Antarctic. Environ.Microbiol. **11:**700-714.
11. **Furuno, S., K. Pazolt, C. Rabe, T. R. Neu, H. Harms, and L. Y. Wick.** 2010. Fungal mycelia allow chemotactic dispersal of polycyclic aromatic hydrocarbon-degrading bacteria in water-unsaturated systems. Environ Microbiol **12:**1391-1398.
12. **Fulthorpe, R. R., L. F. Roesch, A. Riva, and E. W. Triplett.** 2008. Distantly sampled soils carry few species in common. ISME J. **2:**901-910.
13. **Gomes, N. C., L. R. Borges, R. Paranhos, F. N. Pinto, E. Krögerrecklenfort, L. C. Mendonca-Hagler, and K. Smalla.** 2007. Diversity of *ndo* genes in mangrove sediments exposed to different sources of polycyclic aromatic hydrocarbon pollution. Appl. Environ. Microbiol. **73:**7392-7399.
14. **Gomes, N. C., I. A. Kosheleva, W. R. Abraham, and K. Smalla.** 2005. Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. FEMS Microbiol. Ecol. **54:**21-33.
15. **Gomes, N. C. M., L. R. Borges, R. Paranhos, F. N. Pinto, L. C. S. Mendonca-Hagler, and K. Smalla.** 2008. Exploring the diversity of bacterial

communities in sediments of urban mangrove forests. FEMS Microbiol. Ecol. **66:**96-109.

16. **Gomes, N. C. M., H. Heuer, J. Schönfeld, R. Costa, L. Mendonca-Hagler, and K. Smalla.** 2001. Bacterial diversity of the rhizosphere of maize (*Zea mays*) grown in tropical soil studied by temperature gradient gel electrophoresis. Plant Soil **232:**167-180.
17. **Gouy, M., S. Guindon, and O. Gascuel.** 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. **27:**221-224.
18. **Guo, C. L., H. W. Zhou, Y. S. Wong, and N. F. Tam.** 2005. Isolation of PAH-degrading bacteria from mangrove sediments and their biodegradation potential. Mar. Pollut. Bull. **51:**1054-1061.
19. **Habe, H., and T. Omori.** 2003. Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. Biosci. Biotechnol. Biochem. **67:**225-243.
20. **Han, M. V., and C. M. Zmasek.** 2009. phyloXML: XML for evolutionary biology and comparative genomics. BMC Bioinf. **10:**356.
21. **Heuer, H., M. Krsek, P. Baker, K. Smalla, and E. M. Wellington.** 1997. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl. Environ. Microbiol. **63:**3233-3241.
22. **Heuer, H., J. Wieland, A. Schönwälder, N. C. M. Gomes, and K. Smalla.** 2001. Bacterial community profiling using DGGE or TGGE analysis, p. 177-190. *In* I. P. Rouchelle (ed.), Environmental molecular microbiology:protocols and applications. Horizon Scientific Press, Wymondham,United Kingdom.
23. **Hong, J. W., J. Y. Park, and G. M. Gadd.** 2010. Pyrene degradation and copper and zinc uptake by *Fusarium solani* and *Hypocrea lixii* isolated from petrol station soil. J. Appl. Microbiol. **108:**2030-2040.
24. **Hothorn, T., F. Bretz, and P. Westfall.** 2008. Simultaneous inference in general parametric models. Biom. J. **50:**346-363.
25. **Johnsen, A. R., L. Y. Wick, and H. Harms.** 2005. Principles of microbial PAH-degradation in soil. Environ. Pollut. **133:**71-84.

26. **Kanaly, R. A., and S. Harayama.** 2000. Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. J.Bacteriol. **182:**2059-2067.

27. **Kim, S. J., O. Kweon, R. C. Jones, R. D. Edmondson, and C. E. Cerniglia.** 2008. Genomic analysis of polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. Biodegradation **19:**859-881.

28. **Kobayashi, T., Y. Murai, K. Tatsumi, and Y. Iimura.** 2009. Biodegradation of polycyclic aromatic hydrocarbons by Sphingomonas sp. Enhanced by water-extractable organic matter from manure compost. Sci. Total. Environ. **407:**5805-5810.

29. **Kohlmeier, S., T. H. Smits, R. M. Ford, C. Keel, H. Harms, and L. Y. Wick.** 2005. Taking the fungal highway: mobilization of pollutant-degrading bacteria by fungi. Environ. Sci. Technol. **39:**4640-4646.

30. **Kropf, S., H. Heuer, M. Grüning, and K. Smalla.** 2004. Significance test for comparing complex microbial community fingerprints using pairwise similarity measures. J. Microbiol. Methods **57:**187-195.

31. **Leys, N. M., A. Ryngaert, L. Bastiaens, W. Verstraete, E. M. Top, and D. Springael.** 2004. Occurrence and phylogenetic diversity of *Sphingomonas* strains in soils contaminated with polycyclic aromatic hydrocarbons. Appl. Environ. Microbiol. **70:**1944-1955.

32. **Lloyd-Jones, G., A. D. Laurie, D. W. F. Hunter, and R. Fraser.** 1999. Analysis of catabolic genes for naphthalene and phenanthrene degradation in contaminated New Zealand soils. FEMS Microbiol. Ecol. **29:**69-79.

33. **Mattes, T. E., A. K. Alexander, P. M. Richardson, A. C. Munk, C. S. Han, P. Stothard, and N. V. Coleman.** 2008. The genome of *Polaromonas* sp. Strain JS666: insights into the evolution of a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to biotechnology. Appl. Environ. Microbiol. **74:**6405-6416.

34. **Mollea, C., F. Bosco, and B. Ruggeri.** 2005. Fungal biodegradation of naphthalene: microcosms studies. Chemosphere **60:**636-643.

35. **Nayak, A. S., M. H. Vijaykumar, and T. B. Karegoudar.** 2009. Characterization of biosurfactant produced by *Pseudoxanthomonas* sp PNK-04 and its application in bioremediation. Int. Biodeter. Biodegr. **63:**73-79.

36. **Patel, H., A. Gupte, and S. Gupte.** 2009. Biodegradation of Fluoranthene by Basidiomycetes Fungal Isolate Pleurotus Ostreatus HP-1. Appl. Biochem. Biotech. **157:**367-376.

37. **Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner.** 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. **35:**7188-7196.

38. **Schloss, P. D.** 2010. The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. PloS Comput. Biol. **6:**e1000844.

39. **Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber.** 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. **75:**7537-7541.

40. **Sota, M., H. Yano, A. Ono, R. Miyazaki, H. Ishii, H. Genka, E. M. Top, and M. Tsuda.** 2006. Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase. J. Bacteriol. **188:**4057-4067.

41. **Uyttebroek, M., P. Breugelmans, M. Janssen, P. Wattiau, B. Joffe, U. Karlson, J. J. Ortega-Calvo, L. Bastiaens, A. Ryngaert, M. Hausner, and D. Springael.** 2006. Distribution of the *Mycobacterium* community and polycyclic aromatic hydrocarbons (PAHs) among different size fractions of a long-term PAH-contaminated soil. Environ. Microbiol. **8:**836-847.

42. **Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole.** 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. **73:**5261-5267.

43. **Wang, Y., and P. Y. Qian.** 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PloS One **4:**e7401.

44. **Weinert, N., R. Meincke, C. Gottwald, H. Heuer, N. C. Gomes, M. Schloter, G. Berg, and K. Smalla.** 2009. Rhizosphere communities of genetically modified zeaxanthin-accumulating potato plants and their parent cultivar differ

less than those of different potato cultivars. Appl. Environ. Microbiol. **75:**3859-3865.

45. **Yagi, J. M., and E. L. Madsen.** 2009. Diversity, abundance, and consistency of microbial oxygenase expression and biodegradation in a shallow contaminated aquifer. Appl. Environ. Microbiol. **75:**6478-6487.

46. **Yagi, J. M., D. Sims, T. Brettin, D. Bruce, and E. L. Madsen.** 2009. The genome of *Polaromonas* naphthalenivorans strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. Environ. Microbiol. **11:**2253-2270.

47. **Yrjala, K., A. K. Keskinen, M. L. Akerman, C. Fortelius, and T. P. Sipila.** 2010. The rhizosphere and PAH amendment mediate impacts on functional and structural bacterial diversity in sandy peat soil. Environ. Pollut. **158:**1680-1688.

48. **Zeinali, M., M. Vossoughi, and S. K. Ardestani.** 2007. Characterization of a moderate thermophilic Nocardia species able to grow on polycyclic aromatic hydrocarbons. Lett. Appl. Microbiol. **45:**622-628.

49. **Zhang, J., R. Yin, X. G. Lin, W. W. Liu, R. R. Chen, and X. Z. Li.** 2010. Interactive Effect of Biosurfactant and Microorganism to Enhance Phytoremediation for Removal of Aged Polycyclic Aromatic Hydrocarbons from Contaminated Soils. J. Health. Sci. **56:**257-266.

TABLE 5. Differences between microbial DGGE fingerprints for different treatments at Day 21 and 63.

		Bacteria	*Alpha-*	*Beta-*	*Actino-*	*gacA*	*Fungi*
T21	LP-LA	0.0008	0.8472	0.0005	0.0952	0.2768	0.3158
	CP-CA	0.002	0.3084	0.0045	0.5427	0.0109	0.5334
	Phenanthrene	0.0002	0.5551	0.005	0.07	0.004	0.1756
	Soil	0	0.2157	NA	0	0	0
T63	LP-LA	0	0.0021	0.0001	0.0005	0.4539	0.0811
	CP-CA	0.0013	0.0004	0	0	0.0226	0.0457
	Phenanthrene	0	0	0	0	0.0203	0.0362
	Soil	0	0	0	0	0	0

Note: LP: phenanthrene spiked Luvisol; LA: control luvisol; CP: phenanthrene spiked Cambisol; CA: control Cambisol. NA: not analyzed.

TABLE 6. Number of taxa detected.

	T63CA(9816)[a]	T63CP(5084)	T63LA(13261)	T63LP(5893)	Total
phylum	18 (90.8) [b]	13 (94.7)	19 (89.7)	14 (93.3)	21 (91.6)
class	41 (85.9)	38 (91.6)	51 (84.9)	40 (90.9)	51 (87.5)
order	47 (60.3)	43 (76.9)	52 (57.2)	46 (76.5)	57 (64.6)
family	89 (50.8)	79 (72)	100 (48.6)	82 (71)	111(56.8)
genus	193(48.7)	157(60.6)	210 (49.4)	160 (58.1)	271 (52.5)

Note: a: number of sequences detected for the sample; b: percent of sequences which could be classified.

TABLE 7. Taxa enriched in the polluted soils or with significant different relative abundance (%) between two soils

Phylum	Class	Order	Family	Genus	T63CA (3272±640)[a]	T63CP (1695±2079)	T63LA (4420±562)	T63LP (1964±1869)
Proteobacteria[b]	***Alphaproteobacteria***	***Sphingomonadales***	***Sphingomonadaceae***	***Novosphingobium***	0±0	1±0 ***	**0.3±0**[c]	0.8±0 ***
				Sphingomonas	3.4±1	23.4±4 ***	**4±1**	18±5 ***
				Sphingosinicella	0±0	0.1±0	0±0	0.1±0 *
		Caulobacterales	***Caulobacteraceae***	*Caulobacter*	0.07±0	0.1±0 *	0±0	0.1±0
				Phenylobacterium	0.3±0	1.2±1 ***	0.3±0	0.9±0 ***
		Rhizobiales	*Bradyrhizobiaceae*	*Bosea*	0±0	0.2±0 ***	0±0	0.1±0
				Afipia	0±0	0.2±0 **	0±0	0.4±0 ***
	Betaproteobacteria	***Burkholderiales***	***Comamonadaceae***	***Polaromonas***	**0.2±0**	4.1±2 ***	0±0	6.7±3 ***
			Burkholderiaceae	***Burkholderia***	0±0	0.1±0 *	0±0	0.2±0 *
			Oxalobacteraceae	*Duganella*	0±0	0.1±0 *	0.1±0	0.2±0
				Herbaspirillum	0±0	0±0	0±0	0.4±0 ***
				Rhodanobacter	0±0	0.3±0 ***	0±0	1.1±0 ***
	Gammaproteobacteria	*Xanthomonadales*	*Xanthomonadaceae*	*Dokdonella*	0.6±0	0.5±0	0.6±0	1.8±1 ***
				Dyella	0±0	0±0	0.1±0	0.2±0 *
				Pseudoxanthomonas	0±0	0.4±1 *	0±0	0±0
		Pseudomonadales	*Pseudomonadaceae*	*Pseudomonas*	0.1±0	0.6±1 ***	0.1±0	0.3±0
		Legionellales	*Coxiellaceae*	*Aquicella*	0.1±0	0.3±0	0.1±0	0.3±0 ***
			Legionellaceae	*Legionella*	0±0	0±0	0±0	0.3±0 ***
	Deltaproteobacteria	***Bdellovibrionales***	***Bacteriovoracaceae***	***Peredibacter***	0.1±0	0.3±1 ***	0±0	0.2±0 ***
		Nannocystineae	*Nannocystaceae*	*Nannocystis*	0±0	0±0	0±0	0.1±0 **
		Sorangiineae	*Polyangiaceae*	*Byssovorax*	0.1±0	0±0	0.1±0	1±1 ***
		Desulfuromonadales			0.1±0	0±0	**0.6±0**	0.5±0
Actinobacteria	*Actinobacteridae*	***Corynebacterineae***	***Mycobacteriaceae***	***Mycobacterium***	0.1±0	0.2±0 **	0.1±0	0.5±0 **
		Micromonosporineae	*Micromonosporaceae*	*Dactylosporangium*	0±0	0±0	0±0	0.2±0 ***
		Frankineae			**1.1±0**	0.4±0	0.7±0	0.1±0
		Micrococcineae			**3±1**	2.9±1	1.9±0	1.2±1
		Streptomycineae			**1.2±0**	0.7±0	0.8±0	0.3±0
	Acidimicrobidae				1.3±0	0.9±0	**1.9±0**	0.5±0
	Rubrobacteridae				**7.8±1**	3.3±1	4.4±1	2.3±1
Bacteroidetes	*Sphingobacteria*	*Sphingobacteriales*	*Sphingobacteriaceae*	***Mucilaginibacter***	0±0	0.2±0 ***	0±0	0.3±0 ***
Acidobacteria					12.5±0	6.1±2	**14.6±0**	6.8±1
Gemmatimonadetes	*Gemmatimonadetes*	*Gemmatimonadales*	*Gemmatimonadaceae*	*Gemmatimonas*	**4.5±0**	1.2±1	3.8±0	2.2±1

Note: a: average of number sequence for one treatment ±standard deviation; significant higher relative abundance in the polluted soils (*:0.01<p<0.05; **: 0.001<p<0.01; ***: p<0.001). b: bold text, significant enriched in both polluted soils. C: bold numbers, significant higher relative abundance in the soil type.T63: 63 days after phenanthrene spiking; CA: cambisol control; CP: phenanthrene-polluted Cambisol; LA: Luvisol control; LP: phenanthrene-polluted Luvisol.

TABLE 8. Number of OTUs enriched in both soils

Phylum	Cambisol	Luvisol	shared	Total
Acidobacteria	1	0	0	1
Actinobacteria	3	2	1	4
Proteobacteria	44	66	29	81
Bacteroidetes	1	0	0	1
Total	49	69	30	88

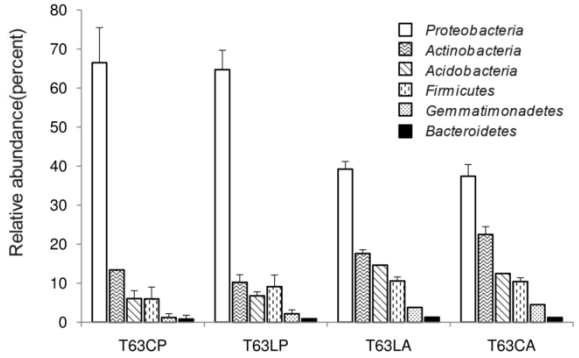

Fig. 1: Relative abundance of dominant phyla detected for different soil samples.

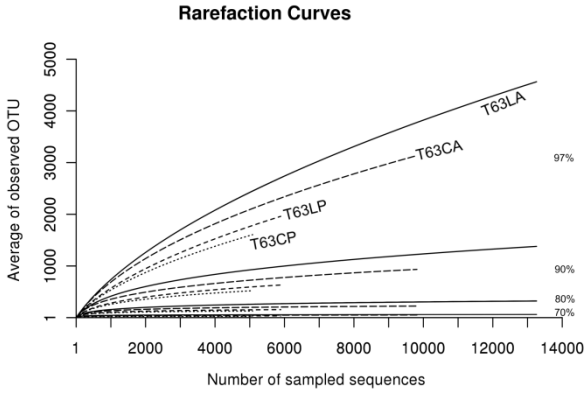

Fig. 2: Plot of rarefaction curves at different similarities.

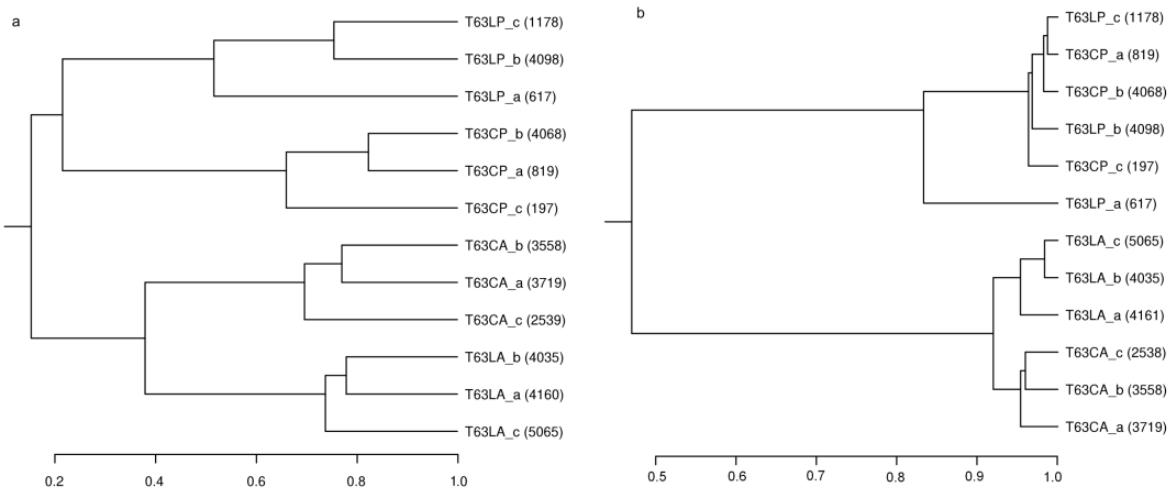

Fig. 3: Furthest-Neighbor-joining cluster of bacterial community structure based on Pearson correlation distance using OTU (a: 97%, and b: 80%) as the species data. Number in brackets: number of sequences per sample.

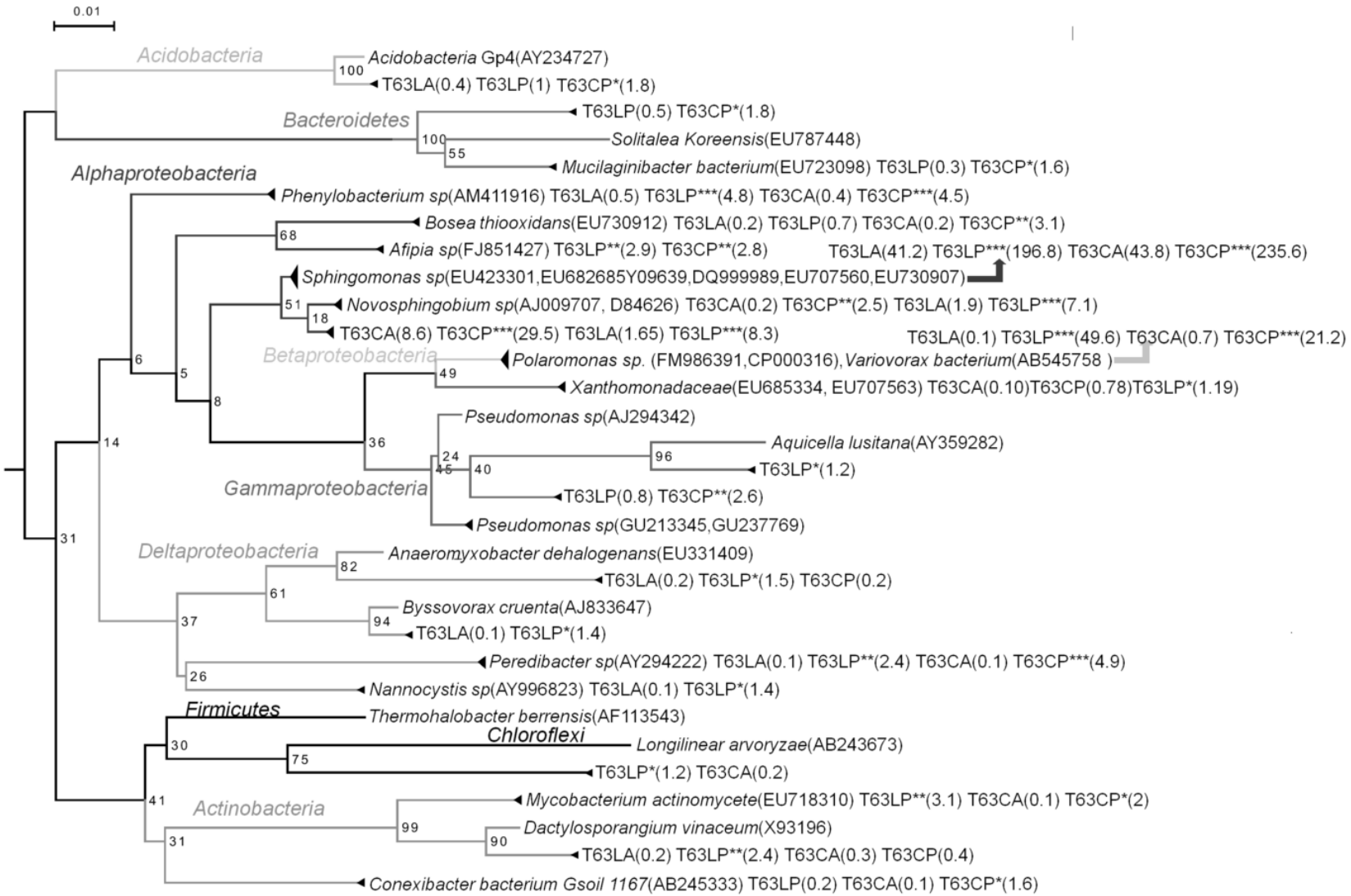

Fig. 4: Neighbor-joining rooted phylogenetic tree based on multiple alignments of representative sequences for significant enriched OTUs (>97% at the species level). Values at each node = (bootstrap value/100) x 100. Number in brackets: Relative abundance in ‰. Significantly higher relative abundance in the polluted soils (*:0.01<p<0.05; **: 0.001<p<0.01; ***: p<0.001). Reference sequences' accesion numbers were also in brackets.

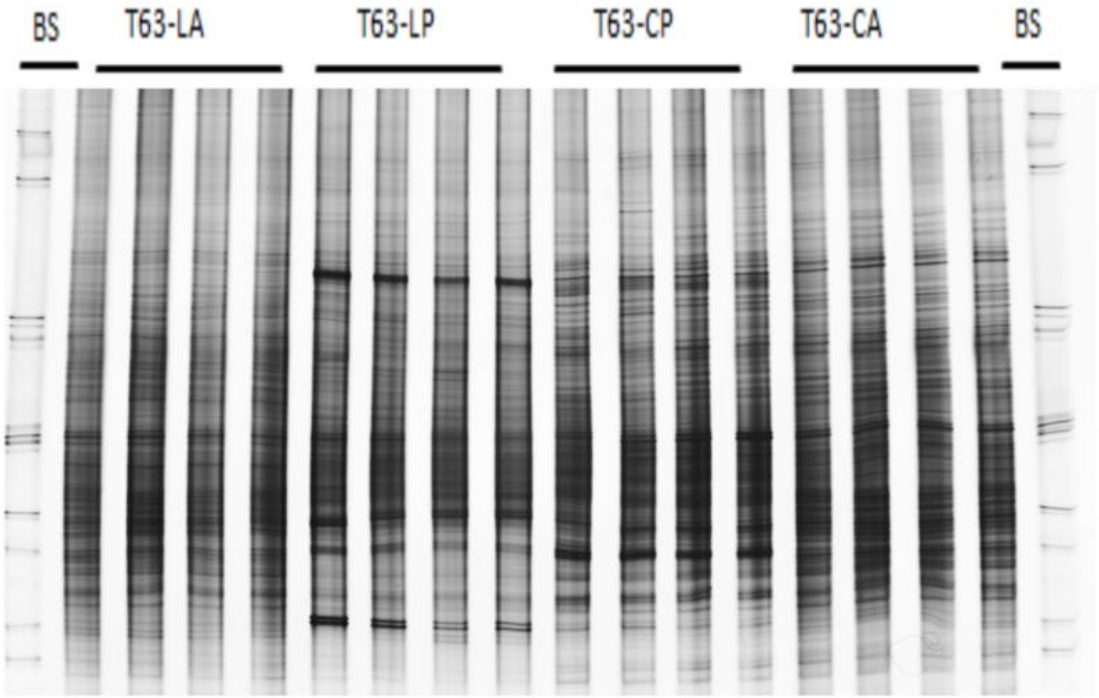

Fig.S1a

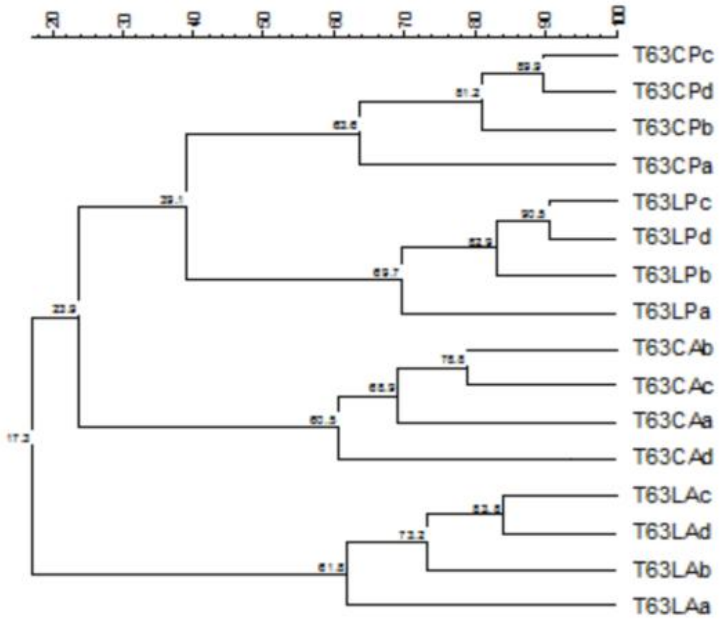

Fig.S1b

Fig. S1: Bacterial DGGE fingerprints and the corresponding UPGMA clusters for soil spiked with phenanthrene (LP, CP) or not (LA, CA) at day 63. BS: bacterial standard consisting of 11 bacterial 16S *rRNA* gene fragment.

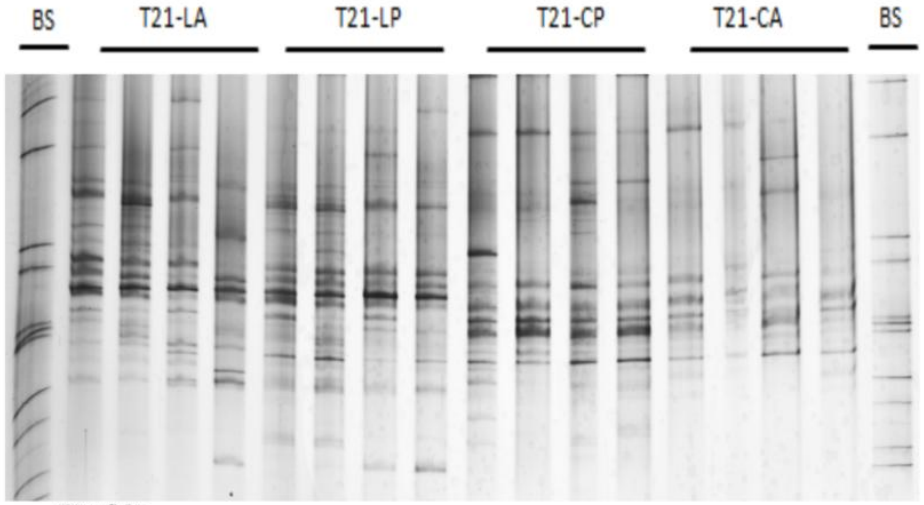

Fig.S2a

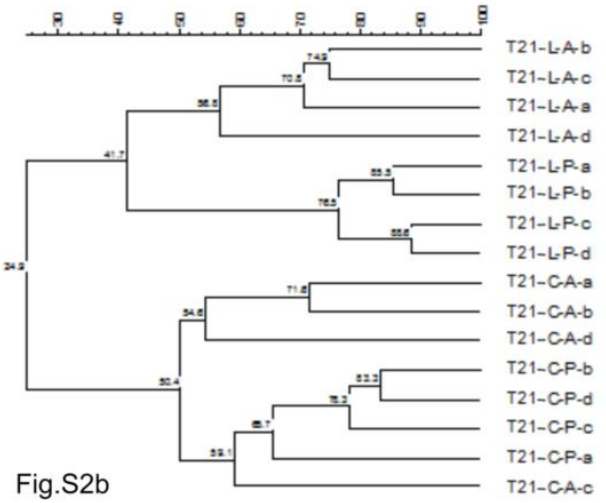

Fig.S2b

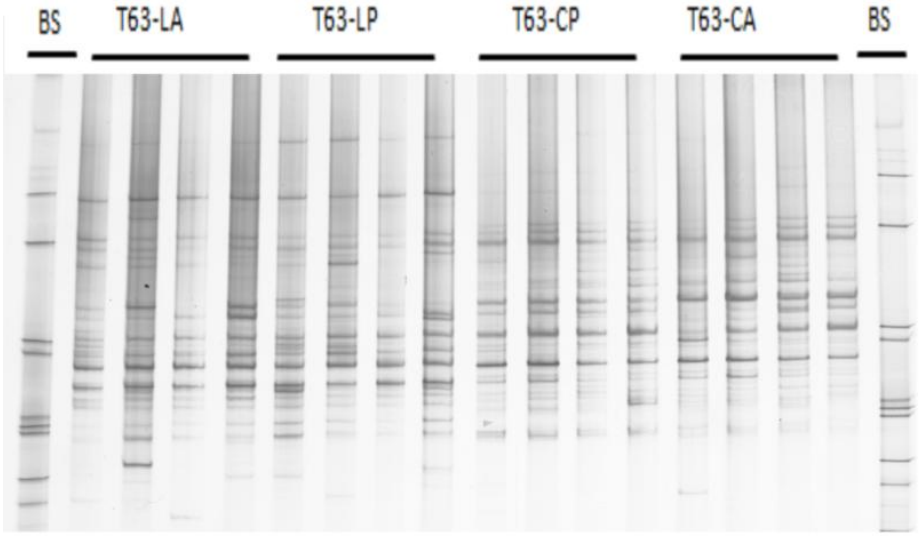

Fig.S2c

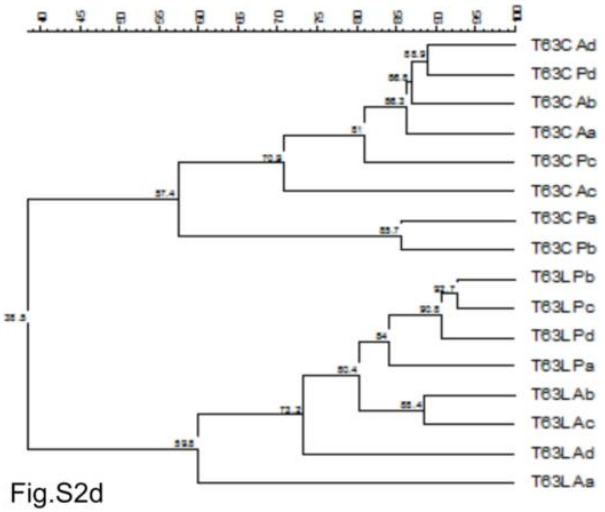

Fig.S2d

Fig. S2: Pseudomonas-specific *gacA* DGGE fingerprints and the corresponding UPGMA clusters for soil spiked with phenanthrene (LP, CP) or not (LA, CA) at days 21 and 63. BS: bacterial standard consisting of 11 bacterial 16S *rRNA* gene fragment.

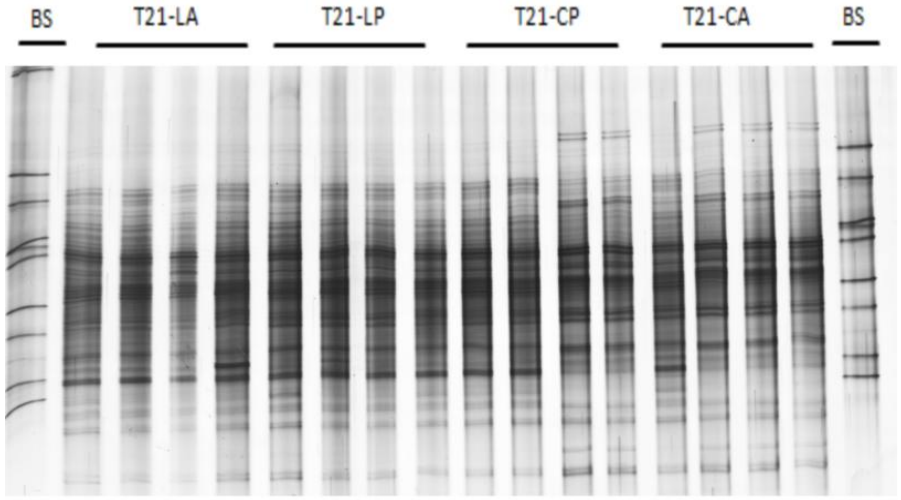

Fig.S3a

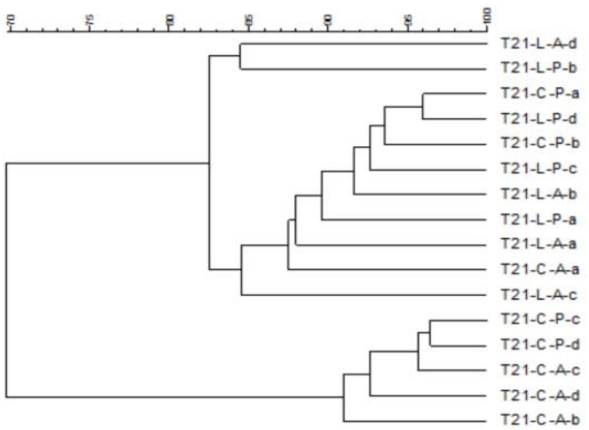

Fig.S3b

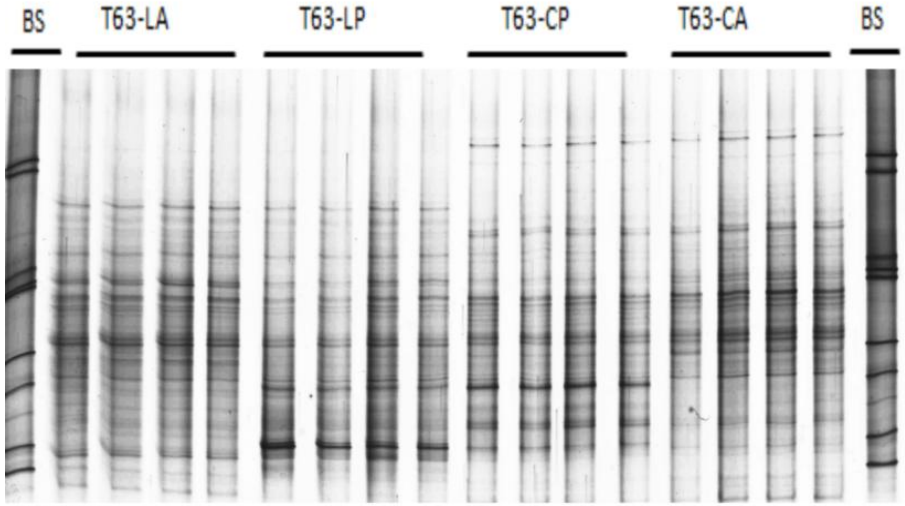

Fig.S3c

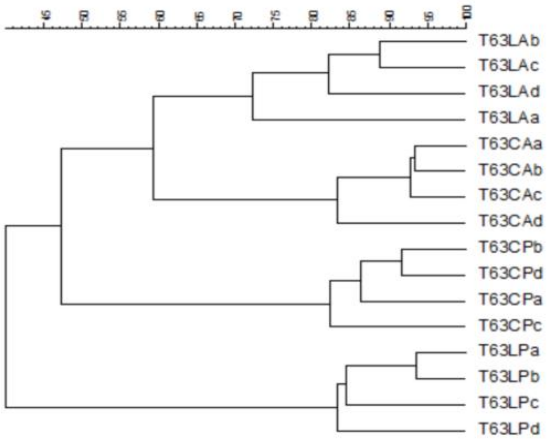

Fig.S3d

Fig. S3: *Actinobacterial*-specific DGGE fingerprints and the corresponding UPGMA clusters for soil spiked with phenanthrene (LP, CP) or not (LA, CA) at days 21 and 63. BS: bacterial standard consisting of 11 bacterial 16S *rRNA* gene fragment.

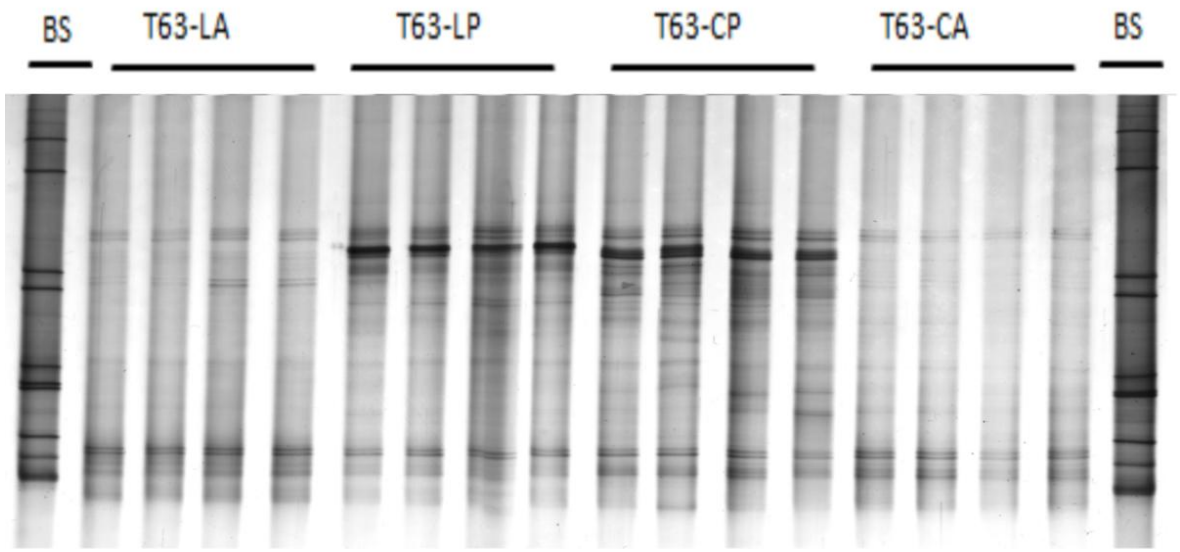

Fig.S4a

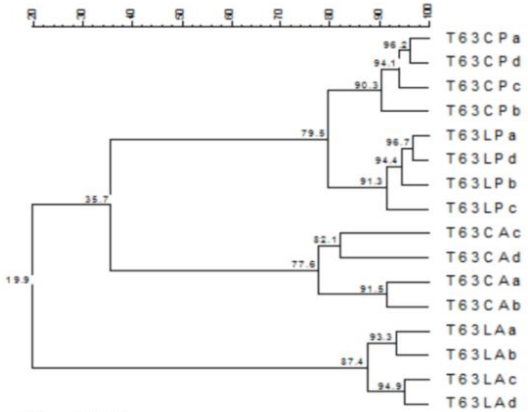

Fig.S4b

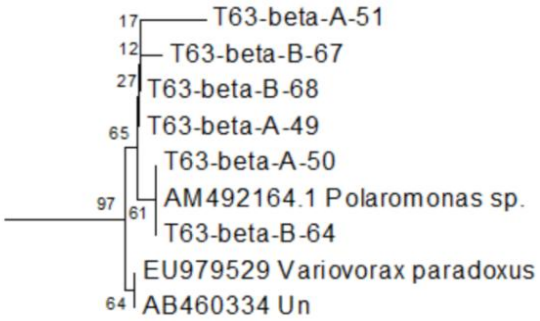

Fig.S4c

Fig . S4: *Betaproteobacteria*-specific DGGE fingerprints and the corresponding UPGMA clusters for soils spiked with phenanthrene (LP, CP) or not (LA, CA) and Neighbor-joining rooted phylogenetic tree based on sequences behind dominant enriched bands in both polluted soils at day 63.

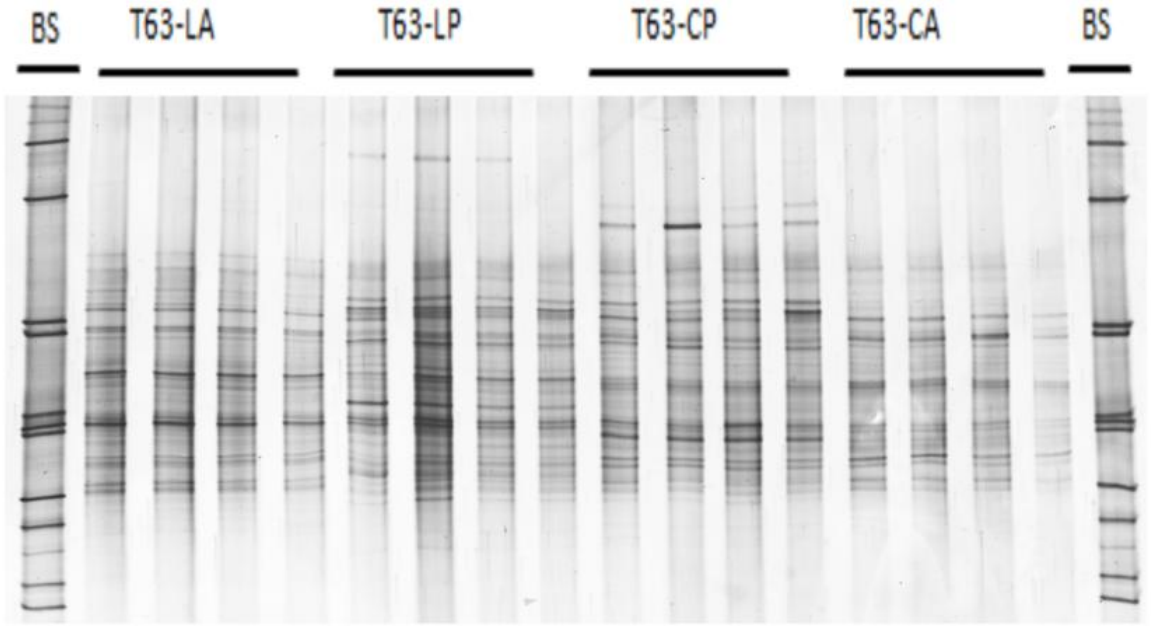

Fig.S5a

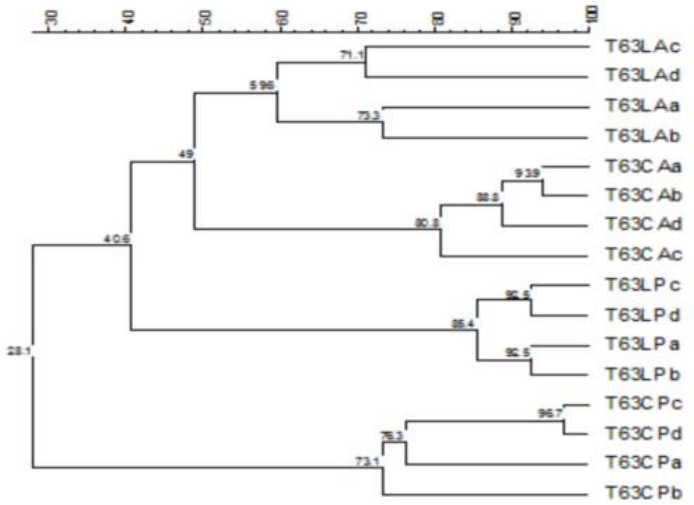

Fig.S5b

Fig. S5: *Alphaproteobacteria*-specific DGGE fingerprints and the corresponding UPGMA clusters for soil spiked with phenanthrene (LP, CP) or not (LA, CA) at day 63.

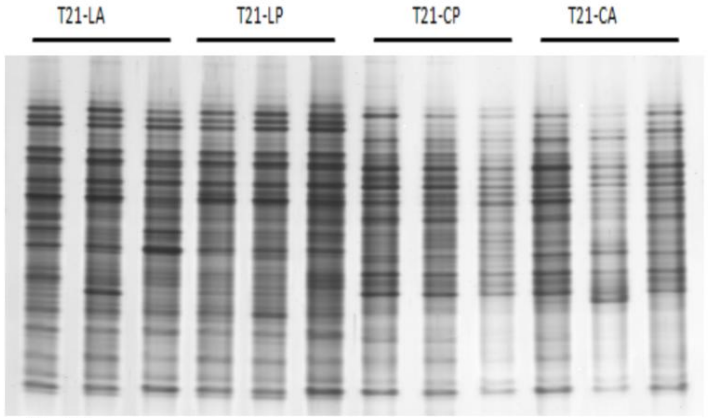

Fig.S6a

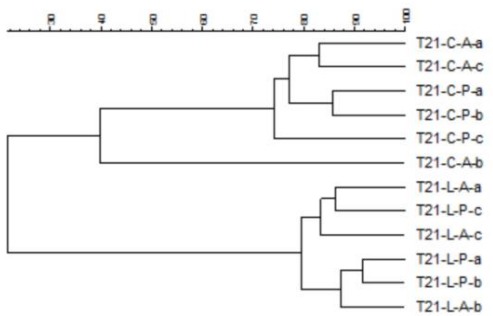

Fig.S6b

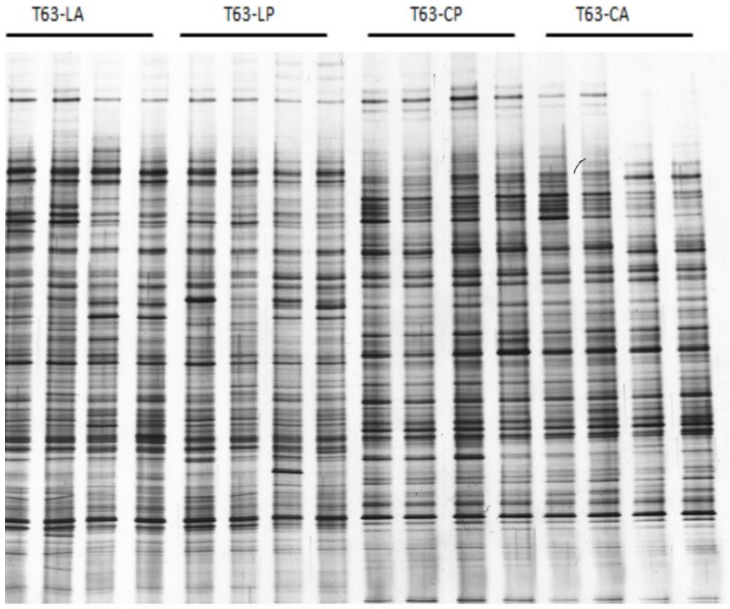

Fig.S6c

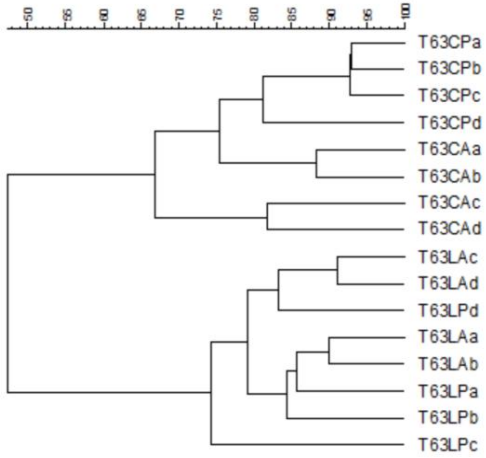

Fig.S6d

Fig. S6: *ITS* DGGE fingerprints for *Fungi* and the corresponding UPGMA clusters for soil spiked with phenanthrene (LP, CP) or not (LA, CA) at days 21 and 63.

Chapter 7: Mineral composition and charcoal determine the bacterial community structure assembled in artificial soils

Guo-Chun Ding, [1] Geertje Johanna Pronk,[2] Holger Heuer,[1] Katja Heister,[2] Ingrid Kögel-Knabner,[2] and Kornelia Smalla[1]*

Julius Kühn-Institut (JKI)—Federal Research Centre for Cultivated Plants, Messeweg 11-12, 38104 Braunschweig, Germany[1];

Lehrstuhl für Bodenkunde,Technische Universität München, 85350 Freising-Weihenstephan, Germany[2]

To study the influence of the mineral composition and charcoal on the establishment of soil bacterial communities, batch experiments were performed with seven artificial soils differing in their of quartz sand montmorillonite, illite ferrihydrite, boehmite and charcoal treated with the same soil microbial inoculant extracted from a Cambisol and with sterile manure as a carbon source. Samples were taken at 1, 9, 31 and 90 days after inoculation. Community changes were monitored by DGGE analysis of bacteria and taxon-specific 16S rRNA gene amplicons from total community DNA. Charcoal artificial soils showed strongest influence on bacterial communities, while the effects of montmorillonite, illite, ferrihydrite and boehmite could only be observed at day 90. Quantitative real time PCR did not reveal significant differences in 16S rRNA gene copy number between the four artificial soils at day 90, while pyrosequencing analysis of amplified bacterial 16Sr RNA gene V3-V4 region showed a distinct community composition and confirmed the strongest influence of charcoal on the bacterial community. Several discriminative taxa between the artificial soils were identified. In summary, this study showed that charcoal and mineral composition shaped the bacterial community structure established in artificial soils and thus might be important factors controlling the characteristics of biogeochemical interfaces formed in soils.

Introduction

Microbes in soil play important roles in maintaining ecosystem services such as the cycling of material and energy, cleaning up polluted sites and enhancing plant growth (Child et al 2007, Peng et al 2010, Yrjala et al 2010). Several efforts have been made to understand the influence of soil characteristics on bacterial community structure and diversity. Physico-chemical factors such as pH, salinity or pore connectivity were reported as major drivers shaping bacterial communities. Fierer and Jackson (2006) found that soil pH values were nicely correlated with detected soil bacterial diversity and richness (Shannon diversity index) (Fierer and Jackson 2006). Based on 21,752 16S RNA gene sequences for 202 environmental samples, Lozupone and Knight (2007) suggested that salinity rather than extremes of temperature, pH or other physico-chemical factors analyzed was the major determinant on microbial community composition. Carson et al (2010) showed that the pore connectivity negative related with active bacterial diversity in soils (Carson et al 2010). All these studies deepened our understanding of the factors shaping dominant bacterial populations in soil.

Soils are mixtures composed of solid, liquid and gaseous phase (Totsche et al 2010). Different chemical and physical gradients, minerals form dynamic interfaces, which are believed to be important for the soil microbial activities and structure (Totsche et al 2010). Unraveling the microbial activities and structure at these biogeochemical interfaces could be a key for a better understanding of the complex soil-microbe interaction. The formation of these interfaces is considered to be affected by the type of particle surfaces present and the assemblage of organic matter and mineral particles with different physio-chemical characteristics forming a large variety of microenvironments that can be colonized by microbes (Huang et al 2005, Totsche et al 2010, Young and Crawford 2004). Thus, soils with different mineral composition provide surfaces with different physico-chemical properties which might influence the characteristics of biogeochemical interfaces in soil formed. But so far, it is still unclear whether the mineral composition of soils influences their microbial community structure. Direct comparison of microbial communities in natural soils with different mineral composition are inherently problematic as (1) different micro-climate conditions, (2) different initial microbial community structure and (3) influxes or efflux of microbes or nutrients from surrounding environments and (4) different soil organic matter content prevail.

To study the influence of soil composition on the establishment of bacterial community structure, artificial soils with different mineral compositions were inoculated with the same microbial consortium, autoclaved manure and incubated in the dark at constant temperature and moisture content. Microbial communities were analyzed by DGGE and amplicon sequencing of 16S rRNA gene fragments amplified from total community DNA (TC-DNA). Thus we could demonstrate that the development of microbial community is indeed depending on the mineral composition of the soil and the presence or absence of charcoal.

Material and methods

Artificial soil incubation

Seven artificial soils each composed of quartz sand (Q) containing a different composition of montmorillonite (M), illite (I), ferrihydrite (F), boehmite (A) and charcoal (C) (Table 1), were inoculated with the microbial fraction extracted from a Cambisol collected at Ultuna (60°N, 17°E) in central Sweden and sterile manure (4.5% mass corresponding to 1.5% organic carbon) as carbon source and incubated under identical conditions. The mixtures were composed so that all compositions had a similar texture consisting of approximately 42 % sand (>63 µm) 52 % silt (6.3-63 µm) and 6% fine silt and clay (<6.3 µm). Microcosms were incubated at room temperature (20±5 °C) in the dark under identical conditions. Samples were taken at days (T) 1, 9, 31 and 90 from the four soil microcosms per treatment and kept at -20 °C before DNA extraction with Bio-101 DNA spin kit for soil (QBiogene, Heidelberg, Germany). Inoculant of microbial fraction was prepared as follows: Fresh soil material (sieved <2 mm), aquarium gravel and demineralized water were shaken for 2 hours. The suspension was then centrifuged at 1,000 *g* for 12 min and the supernatant was transferred to a new vial and centrifuged at 3,470 *g* for 30 min. The pellet from the second centrifugation step was then re-suspended with de-mineralized water and 15 ml inoculant solution was obtained from one gram of soil. Sixty ml inoculant solution was used per kg of artificial soil. The pH, C and N content of the artificial soils was determined at day 1 and day 90 of incubation (Table S2). The pH of the samples was determined in duplicate in 0.01 M $CaCl_2$ at a soil: solution ratio of 1:2.5. Carbon and nitrogen content were determined in duplicate by combustion and chromatographic separation with an elemental analyzer (Hekatech, Euro EA 3000, Wegberg,

Germany). All analyses were performed in duplicate for each of the three replicates present for each soil composition.

DGGE analysis of 16S rRNA gene amplicon

Bacterial 16S rRNA gene fragments from the soil samples were amplified with the primers F984-GC and R1378 as described by Heuer et al (1997). A nested-PCR approach was also applied for amplification of 16S rRNA genes of *Actinobacteria* and *Proteobacteria* specific groups as previously described (Costa et al 2007, Gomes et al 2001, Heuer et al 1997). For all nested PCR, the first amplication were of 30 thermal cycles and then the amplicons were checked on agarose gel to make proper dilution of templates for the second amplification with 25 thermal cycles. DGGE of the 16S rRNA gene amplicons were performed according to Gomes et al (2005). The DGGE gels were silver stained according Heuer et al (2001). Software package GelCompare 4.5 was used for cluster analysis of DGGE profiles. Dendrograms were constructed based pair wise Pearson correlation index within a gel using unweighted pair group method using arithmetic averages (UPGMA).

Determination of 16S rRNA gene copy number

Primers and a TaqMan probe to quantify 16S rRNA genes by qPCR were previously described (Takai and Horikoshi 2000). PCR amplifications were performed in a 50 µl reaction volume containing 1.25 U TrueStart polymerase (Fermentas, St. Leon-Rot, Germany), 0.2 mM each deoxynucleoside triphosphate, 2.5 mM $MgCl_2$, and 0.25 µM primers and probe. Thermocycles were 10 min at 94°C and 40 cycles consisting of 15 s at 95°C, 15 s at 50°C, and 60 s at 60°C. Templates to generate standard curves were prepared by serial dilutions of gel-purified PCR products from *Escherichia coli* 16S rRNA genes.

Pyrosequencing and sequences analysis

Samples (QM, QI, QMC and QIF in Table 1) collected at Day 90 were further studied by pyrosequencing with three replicates per treatment. PCR amplification of the V3-V4 region and following sequencing were done at Biotechnology Innovation Center (BIOCANT) in Portugal.

Pyrosequencing data was evaluated as previously described. Briefly, noise data were removed based on standard alone BLASTN reports. Sequences with a length above 200 bp were subjected to further analyses. Software package mothur (v1.14.0) was used for multiple alignment, OTU assignment and rarefaction analysis (Schloss et al 2009, Schloss 2010). Sequences were further classified by the software RDP MultiClassifier at >80% confidence (Wang et al 2007). Aligned sequences, their corresponding taxonomy, as well as OTU assignment were stored in a local MYSQL database. OTU reports were generated with a perl script. Data summary at different taxonomic levels, community structure (Pearson index distance) and diversity (Jaccard index) analysis, diversity index calculation (simboot), statistic analysis (multcomp) (Hothorn et al 2008), plot of rarefaction curves were done with software R (2.12) with/without add-on packages. Principal component analysis (PCA) was performed based on relative abundance of OTUs belonging to different taxonomic groups (phyla, class, order, family, genus levels) which have more than 15 OTUs. MANOVA with three principal components explaining more than 69% of the total variance was performed to test the influences of artificial soil composition on the community structure according to Glimm et al (1997). Representative sequences for OTUs with discriminative relative abundance were retrieved from the local database. A neighbor joining phylogenetic tree was constructed based on pair wise Jukes & Cantor DNA distances using the software package SeaView4(Gouy et al 2010). Further editions on the phylogenetic tree were done with software Archaeopteryx (Han and Zmasek 2009).

Results

PCR-DGGE analysis of soil bacterial communities assembled in artificial soils

Community fingerprints of *Bacteria, Betaproteobacteria* and *Actinobacteria* were generated and the profiles were clustered by UPGMA based on Pearson correlation indices.

Complex patterns were already observed 9 days after inoculation (data not shown). The presence of charcoal was found to strongly influence the bacterial community structure, as UPGMA analysis of bacterial community profiles revealed that separate clusters were always formed for those samples with charcoal. Effects of the clay minerals montmorillonite and illite on bacterial community patterns were clearly observed at day 90 (Fig. S1). Bacterial community patterns of artificial soils with

montmorillonite (QM) differed from those samples with illite (QI). The betaproteobacterial (Fig. S9) and actinobacterial (Fig. S10) patterns of soils with ferrihydrite or aluminium hydroxide clustered separately from the other patterns, but no clear differences were observed between the ferrihydrite and aluminium hydroxide containing soils. Base on the DGGE analyses four artificial soils (QMC, QM, QI and QIF) were selected for 16S rRNA amplicon sequencing.

Real-time PCR quantification (Q-PCR) of 16S rRNA gene

To estimate the number of bacteria in the artificial soils, 16S rRNA gene copy numbers were quantified by quantitative real-time PCR for samples QMC, QM, QI and QIF. The 16S rRNA gene copies number per gram soil ranged between $1x10^9$ and $3\ x10^{10}$ for most samples except for two samples (T9 QMC-b, T90 QIF-a) with about $2\ x10^6$. The 16S rRNA gene copy numbers were log10 transformed for statistic analysis by one way ANOVA in conjunction with TukeyHSD tests. No significant difference between any pair of treatments was observed ($p<0.05$).

Pyrosequencing analysis

The microbial community of artificial soils QMC, QM, QI and QIF collected at 90 days after inoculation, were further analyzed by barcoded amplicon pyrosequencing of V3-V4 region of 16S rRNA gene. All together 36,164 sequences were obtained for 12 artificial soil samples and 3,535 sequences were obtained for inoculants. A total of 13 phyla were detected in the artificial soils, among them *Firmicutes* (14875), *Proteobacteria* (14124), *Bacteroidetes* (3958) and *Actinobacteria* (1769) are the most frequently detected. In addition, 1357 sequences classified as Bacteria could not be assigned to any known phylum by Naïve Bayer classifier. Although the same number of phyla were also detected from the inoculant with *Proteobacteria* (1984), *Actinobacteria* (682), *Acidobacteria* (228) and *Gemmatinonadetes* (219) as the dominant groups their composition was clearly different from those in the artificial soils. Most remarkable is the high abundance of sequences affiliated to the *Firmicutes* in the artificial soils, while sequences belonging to this phylum were not frequently detected in the inoculant. In contrast, *Acidobacteria* were frequently detected in the inoculant but only rarely detected in the artificial soils. Sequences affiliated to *Cyanobacteria*, *Nitrospira* and WS3 were interestingly only detected in the inoculant. Details of relative abundance of different taxa were summarized (See

Table S1). In total, 8527 OTU (>97% similarity) were detected in all samples. Higher detected diversity was found for the artificial soils QIF, QI and QM than for QMC (Fig. 2). Similar trends were also observed based on OTU assigned at other similarity levels (90%, 80% and 70%). High Shannon, Simpson and Pielou's evenness index were acquired based on the OTU (>97%) reports (Fig. 3). This result suggested that both the diversity of the bacterial community structure in artificial soils and the inoculant and the evenness of OTU with >97% (species level) were high.

The presence of charcoal dramatically influenced the bacterial community established. Although the overlap of OTUs (>97%) between all samples was low (Fig. 4 b), a separated cluster still formed for QMC. Community structure analysis based on Pearson correlation index distance confirmed the strong effects of charcoal on the dominant bacterial community (Fig. 4 a). More than 70% similarities were shared between three replicates of QMC. Other artificial soil samples except the replicates for QM, cluster together. MANOVA with the first three principle components revealed a significant difference in bacterial community structure among the four different artificial soils. To compare the community structure of different taxonomic groups (phylum, class, order, family, genus), multiple PCA in conjunction with MANOVA with the first three principle components were performed. All detected dominant phyla (*Proteobacteria, Firmicutes, Actinobacteria, Bacteroidetes, Gemmatimonadetes*) have significant difference in community structure among treatments (See Table 2).

To further study the influence of the mineral composition and of the presence of charcoal on the relative abundance of different taxa, multiple Tukey's tests under a generalized linear model via logistic function for binomial data were performed (Table 2). Discriminative taxa were identified between artificial soils. The relative abundance of many genera such as *Phenylobacterium*, *Devosia*, *Rhizobium, Sphingomonas*, *Cupriavidus*, *Massilia*, *Luteimonas*, *Pseudoxanthomonas*, *Peredibacter*, *Lysinibacillus*, *Oceanobacillus*, *Ammoniphilus, Cohnella*, *Paenisporosarcina*, *Sporosarcina*, *Turicibacter*, *Streptomyces*, *Chitinophaga*, *Adhaeribacter*, *Pontibacter*, *Sphingobacterium*, *Gemmatimonas* were significant different between QMC and other artificial soils. Among them, *Sphingomonas*, *Lysinibacillus* and *Streptomyces* were dominant groups (>1%) in artificial soil QMC, but not in others. The majority *Streptomyces* sequences in QMC was most similar to *S. coeruleorbidus* (HQ711986) (Fig. S6). Discriminative genera between two artificial soils with different clay minerals were also found. In QI compared to QM, the relative abundances of

Janthinobacterium, *Nitrosospira*, *Bacillus*, *Marinibacillus*, *Oceanobacillus*, *Sporosarcina*, *Arthrobacter, Streptomyces*, *Adhaeribacter* were significantly higher, while that of *Cellvibrio*, *Pseudomonas*, *Luteimonas, Lysobacter* and *Pontibacter* were significantly lower. Interestingly, *Brevundimonas, Cellvibro, Luteimonas, Lysobacter* and *Pontibacter* were found significant enriched in the artificial soil with Ferrihydrite (QIF) compared to QI. At the species level, a total of 265 discriminative OTUs were identified between different artificial soils samples. Most of them belonged to *Proteobacteria* (125), *Firmicutes* (84), *Bacteroidetes* (37) and *Actinobacteria* (13). Representative sequences for these OTUs and their closest related reference sequences were subjected for phylogenetic analyses (Fig. S3-6). At the species level (>97%), some OTU belonging to same genus were preferentially detected in different artificial soils. For example, sequences sharing high similarities with the 16S rRNA gene sequences of *Sphingomonas sp.* (FJ948169, HM484371) were mainly detected for QMC soils, while a big part of those sequences similar to *Sphingomonas sp* (AB512774) were observed in the QM soil (Fig. S3). Also within the genera *Pseudomonas* (Fig. S3), *Bacillus* (Fig. S4) or *Pedobacter* (Fig. S5), an artificial soil type-dependent preferential occurrence of particular species was observed. Interestingly, many species were exclusively detected in QMC soil (Fig. S3-6). These species were most similar to *Sphingomonas* sp. (HM484371), *Parvibaculum* (AY007683), *Devosia* (EF433461, AM393883), *Luteimonas aquatica* (EF626688), *Lysobacter* sp. (EU833988), *Herbaspirillum* sp. (GU377118) (Fig. S3), *Bacillus* sp. (FR667181) (Fig. S4), *Pedobacter* (FJ377315) (Fig. S5), *Myxobacterium* (AF482687, AB246771) and *Gemmatimonas aurantiaca* (HM154525) (Fig. S6).

Discussion

To simulate the formation of biogeochemical interfaces in soil, the influence of different minerals and charcoal on the soil microbial community established was investigated by means of artificial soils. The artificial soils differing in the composition of clay minerals (illit, montmorrilonite), metal oxides (Fe, Al) and charcoal provided virgin-like surfaces with different physico-chemical characteristics to the microbes inoculated. The microbial inoculant was obtained from a fresh Cambisol but likely presents only a fraction of bacteria that are easily to be detached. Although autoclaved manure was added to all artificial soils as a nutrient source, in contrast to real soils complex soil organic matter was lacking. Compared to soils directly

sampled from the field, the advantages of the use of artificial soils to investigate the influence of the mineral composition and of charcoal are: (1) known physico-chemical characteristics of the components, (2) defined soil mineral composition, (3) identical initial microbial community structure of the inoculant, (4) same incubation conditions and (5) exclusion of the influences of influx and efflux of microbes/nutrients from surroundings.

Although all artificial soils shared a high proportion of sand and silt sized quartz (94 %), the rather small amounts of clay minerals, ferrihydrite or aluminium hydroxide or charcoal added triggered the establishment of distinct bacterial communities as revealed by DGGE and pyrosequence analysis of 16S rRNA gene fragments amplified from total community DNA. The diversity indices (Shannon and Simpson) of OTUs detected from all artificial soils were high and their distribution according to the Pielou's evenness index was even. High richness and evenness of bacteria at species levels are typical for soils. In chapter 6, the bacterial community structure in the same Cambisol was analyzed by pyrosequencing of 16S rRNA gene V3-V4 regions. Compared to the Cambisol, the detected richness and evenness of OTUs for artificial soils were lower (Fig. S7). Most likely, artificial soils offer a lower diversity of micro-niches than natural soils. Furthermore, artificial soils have a lower clay content (approx. 6%) than the Cambisol (37%) and clay content was recently shown to be negatively correlated to pore connectivity and positively to bacterial diversity (Carson et al 2010). But it cannot be excluded that lower detected diversity in the artificial soil could be also due to the reduced richness of the inoculant. The dominant bacterial community structure established in the artificial soils was completely different from those in the inoculants and in the Cambisol (Fig. S8). The community structure in the inoculants is more similar to those in the Cambisol than to that in the artificial soil.

Nevertheless, both DGGE and pyrosequencing analyses suggested that bacterial community structure reassembled on artificial soils depends on the mineral composition and charcoal. Separated cluster were formed for artificial soils with different clay minerals or with charcoal based on bacterial profile. In addition, the effects of iron oxides were observed in betaproteobacterial (Fig. S9) and actinobacterial (Fig. S10) profiles. The effect of charcoal and clay minerals on the bacterial community structure was also confirmed by pyrosequencing analysis. In addition to the separated clusters, several taxa were identified with significant differences between the four artificial soils (Table 2). Discriminative taxa belonging to

Betaproteobacteria and *Actinobacteria* could be identified between QI and QIF, but their relative abundance is very low (<1%). Low relative abundance of these discriminative taxa might not be enough to change the total bacterial community patterns based on bacterial DGGE and pyrosequencing. Group-specific primers for *Betaproteobacteria* and *Actinobacteria* improve the resolution and allow studying less abundant populations and thus to improve the resolution level. It is also worth to note that high similarity of bacterial profiles pattern was observed for replicates of artificial soils with the same composition. Pyrosequencing analysis also confirmed this finding. These results do show that mineral composition and charcoal are important drivers of bacterial community structure. Interestingly, in contrast to the community structure which is stable among replicates (except QM-a), overlap of OTUs at species level was very low (<25%) between replicates. The majority of OTUs which are less abundant might only be detected by chance. Therefore, even though high throughput sequencing provides more detailed information on bacterial community structure than we ever had before. Still rare sequences need to be cautiously evaluated. In addition, the species diversity in complex environmental samples such as soils might be not easily estimated based on limited number of sequences.

Compared with illite, montmorillonite, ferrihydrite and boehmite, charcoal was found to be the strongest factor which controlled the bacterial community structure. Discriminative bacterial community structure for artificial soils with charcoal was even observed 9 days after inoculation. Significantly higher relative abundances of several taxonomic groups such as *Devosia*, *Sphingomonas, Lysinibacillus* and *Streptomyces* were found in artificial soils with charcoal. The effect of charcoal on soil microbial activities or abundance of genes involved in nitrogen cycling was also observed in other studies. For example, Ball et al (2010) observed increased abundance of nitrification and ammonium-oxidizing bacteria in charcoal rich natural soils (Ball et al 2010). In this study, the relative abundance of *Nitrosospira* was found significantly higher in QMC than in QM. Interestingly, the relative abundance of OTUs affiliated to *Rhizobium* were significantly higher in QMC than others artificial soils. The presence of charcoal might enhance the abundance of population involved in the nitrogen cycle. The strong adsorption ability of charcoal could also influence the microbial activities. In the study of Rhodes et al.,(Rhodes et al 2010), the mineralization of phenanthrene by the indigenous microbial community was reduced by more than 99%, when only 0.1% activated charcoal was added to in the soils. It was assumed

that the decreasing mineralization rate due to enhanced phenanthrene sorption by charcoal resulted in a reduced bioavailability. In this study the proportion of charcoal was much higher (2%) and thus nutrient rich charcoal particles and oligotrophic surroundings outside charcoal particles could coexist in the QMC soil. Both bacteria with strong competence under oligotrophic conditions and those which are competent in nutrient rich conditions might have an advantage in the artificial soil with charcoal. In this study, the relative abundance of *Sphingomonas* was significantly higher in QMC than in others. Species belonging to this genus were identified in different studies as important degraders of pollutants such as polycyclic aromatic hydrocarbons(Cho and Kim 2001, Leys et al 2004, Vacca et al 2005, Zhou et al 2006) , chlorinated compounds (Leigh et al 2007) and wildly exists in soil, water distribution system, marine and other oligotrophic environments (Johnsen et al 2005). The relative abundance of *Lysinibacillus* and *Streptomyces* were also significantly higher in the QMC than in others artificial soil. Further experiments are needed to identify bacteria dominant in charcoal particles. Interestingly, biochar was used to improve soil in nutrient conditions, the ability to suppress soil-borne phytopathogens. But the microbial mechanisms are rarely known. This study clearly shown that functional groups such as those involved in nitrogen cycling (*Nitrosospira*, *Rhizobium*), or taxa (*Streptomyces*, *Luteimonas*, *Pseduomonas*) with known antagonist were enriched in the artificial soil with charcoal.

In summary, this study first time illustrates that mineral composition and charcoal is driving force in shaping soil bacterial community structure.

ACKNOWLEDGMENTS This work was supported by DFG SPP1315 (SM59/8–1, SP255/19–1, and KO1035/33–1) and the BMBF project MÄQNU.

Reference

Ball PN, MacKenzie MD, DeLuca TH, Holben WE (2010). Wildfire and charcoal enhance nitrification and ammonium-oxidizing bacterial abundance in dry Montane forest soils. *J Environ Qual* **39:** 1243-1253.

Brim H, Heuer H, Krögerrecklenfort E, Mergeay M, Smalla K (1999). Characterization of the bacterial community of a zinc-polluted soil. *Can J Microbiol* **45:** 326-338.

Carson JK, Gonzalez-Quinones V, Murphy DV, Hinz C, Shaw JA, Gleeson DB (2010). Low pore connectivity increases bacterial diversity in soil. *Appl Environ Microbiol* **76:** 3936-3942.

Child R, Miller CD, Liang Y, Sims RC, Anderson AJ (2007). Pyrene mineralization by Mycobacterium sp strain KMS in a Barley rhizosphere. *Journal of Environmental Quality* **36:** 1260-1265.

Cho JC, Kim SJ (2001). Detection of mega plasmid from polycyclic aromatic hydrocarbon-degrading Sphingomonas sp. strain KS14. *J Mol Microbiol Biotechnol* **3:** 503-506.

Costa R, Gomes NC, Krögerrecklenfort E, Opelt K, Berg G, Smalla K (2007). Pseudomonas community structure and antagonistic potential in the rhizosphere: insights gained by combining phylogenetic and functional gene-based analyses. *Environ Microbiol* **9:** 2260-2273.

Fierer N, Jackson RB (2006). The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* **103:** 626-631.

Glimm E, Heuer H, Engelen B, Smalla K, Backhaus H (1997). Statistical comparisons of community catabolic profiles. *J Microbiol Meth* **30:** 71-80.

Gomes NC, Kosheleva IA, Abraham WR, Smalla K (2005). Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. *FEMS MicrobiolEcol* **54:** 21-33.

Gomes NCM, Heuer H, Schonfeld J, Costa R, Mendonca-Hagler L, Smalla K (2001). Bacterial diversity of the rhizosphere of maize (Zea mays) grown in tropical soil studied by temperature gradient gel electrophoresis. *Plant Soil* **232:** 167-180.

Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27:** 221-224.

Han MV, Zmasek CM (2009). phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* **10:** 356.

Heuer H, Krsek M, Baker P, Smalla K, Wellington EM (1997). Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. *Appl Environ Microbiol* **63:** 3233-3241.

Heuer H, Wieland J, Schönwälder A, Gomes NCM, Smalla K (2001). Bacterial community profiling using DGGE or TGGE analysis. In: Rouchelle IP (ed). *Environmental molecular microbiology:protocols and applications*. Horizon Scientific Press: Wymondham,United Kingdom. pp 177-190.

Hothorn T, Bretz F, Westfall P (2008). Simultaneous inference in general parametric models. *Biom J* **50:** 346-363.

Huang PM, Wang MK, Chiu CY (2005). Soil mineral-organic matter-microbe interactions: Impacts on biogeochemical processes and biodiversity in soils. *Pedobiologia* **49:** 609-635.

Johnsen AR, Wick LY, Harms H (2005). Principles of microbial PAH-degradation in soil. *EnvironPollut* **133:** 71-84.

Leigh MB, Pellizari VH, Uhlik O, Sutka R, Rodrigues J, Ostrom NE *et al* (2007). Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). *ISME J* **1:** 134-148.

Leys NM, Ryngaert A, Bastiaens L, Verstraete W, Top EM, Springael D (2004). Occurrence and phylogenetic diversity of Sphingomonas strains in soils contaminated with polycyclic aromatic hydrocarbons. *Appl Environ Microbiol* **70:** 1944-1955.

Lozupone CA, Knight R (2007). Global patterns in bacterial diversity. *Proc Natl Acad Sci U S A* **104:** 11436-11440.

Peng JJ, Cai C, Qiao M, Li H, Zhu YG (2010). Dynamic changes in functional gene copy numbers and microbial communities during degradation of pyrene in soils. *Environ Pollut* **158:** 2872-2879.

Rhodes AH, McAllister LE, Chen R, Semple KT (2010). Impact of activated charcoal on the mineralisation of 14C-phenanthrene in soils. *Chemosphere* **79:** 463-469.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75:** 7537-7541.

Schloss PD (2010). The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS Comput Biol* **6:** e1000844.

Takai K, Horikoshi K (2000). Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *ApplEnvironMicrobiol* **66:** 5066-5072.

Totsche KU, Rennert T, Gerzabek MH, Kögel-Knabner I, Smalla K, Spiteller M *et al* (2010). Biogeochemical interfaces in soil: The interdisciplinary challenge for soil science. *Journal of Plant Nutrition and Soil Science* **173:** 88-99.

Vacca DJ, Bleam WF, Hickey WJ (2005). Isolation of soil bacteria adapted to degrade humic acid-sorbed phenanthrene. *ApplEnvironMicrobiol* **71:** 3797-3805.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73:** 5261-5267.

Young IM, Crawford JW (2004). Interactions and self-organization in the soil-microbe complex. *Science* **304:** 1634-1637.

Yrjala K, Keskinen AK, Akerman ML, Fortelius C, Sipila TP (2010). The rhizosphere and PAH amendment mediate impacts on functional and structural bacterial diversity in sandy peat soil. *Environ Pollut* **158:** 1680-1688.

Zhou HW, Guo CL, Wong YS, Tam NF (2006). Genetic diversity of dioxygenase genes in polycyclic aromatic hydrocarbon-degrading bacteria isolated from mangrove sediments. *FEMS MicrobiolLett* **262:** 148-157.

Table 1 Composition of the artificial soils

	QM	QI	QMI	QMC	QIF	QIA	QIFC
quartz	99	99	99	97	98	98	96
montmorillonite	6	0	3	6	0	0	0
illite	0	6	3	0	6	6	6
ferrihydrite	0	0	0	0	1	0	1
charcoal	0	0	0	2	0	0	2
boehmite	0	0	0	0	0	1	0

Table2: Relative abundance (%) of discrimative taxanomic groups in different artificial soils 90 days after inoculation.

Phylum	Class	Order	Family	Genus	Ino	QI	QIF	QM	QMC
Proteobacteria	***Alphaproteobacteria***	*Caulobacterales***	*Caulobacteraceae***	*Brevundimonas***	0.1	1.1±1 b	**2.1±1 a**	1.2±1 b	**2.4±1 a**
				Phenylobacterium	0.2	0.3±0 b	0.2±0 b	0.3±0 ab	**0.6±0 a**
		Rhizobiales	***Hyphomicrobiaceae***	***Devosia***	0.1	1.3±0 b	0.9±0 c	1±0 bc	**2.4±1 a**
			Rhizobiaceae	*Rhizobium*	0	0.1±0 b	0±0 b	0.1±0 b	**0.5±0 a**
		*Sphingomonadales***	***Sphingomonadaceae***	*Sphingomonas***	12.8	0.4±0 b	0.2±0 b	0.2±0 b	**3.9±1 a**
	*Betaproteobacteria***	*Burkholderiales***	*Burkholderiaceae*	*Cupriavidus*	0	0±0 b	0±0 b	0±0 b	**0.3±0 a**
			*Oxalobacteraceae**	*Janthinobacterium***	0.1	0.8±0 a	0.4±0 b	0±0 c	0±0 c
				Massilia	0.3	0±0 b	0±0 b	0.1±0 b	**0.8±1 a**
		*Nitrosomonadales**	*Nitrosomonadaceae**	*Nitrosospira**	0	**0.7±0 a**	0.3±0 bc	0.2±0 c	0.5±0 ab
	*Gammaproteobacteria***	*Pseudomonadales*	*Pseudomonadaceae*	*Cellvibrio*	0	0±0 c	0.5±0 b	**14.9±26 a**	0.3±0 b
				*Pseudomonas***	0.6	0.1±0 b	0.2±0 b	**1.6±2 a**	**1.5±1 a**
		Xanthomonadales	***Xanthomonadaceae***	***Luteimonas***	0	4.5±1 c	5.6±1 b	6.3±3 b	**8.9±4 a**
				*Lysobacter**	0.9	0.4±0 c	**2.2±2 a**	1.9±1 b	0.5±0 c
				Pseudoxanthomonas	0	0.1±0 b	0.1±0 b	0.1±0 b	**0.4±0 a**
	*Deltaproteobacteria***	*Bdellovibrionales**	*Bacteriovoracaceae*	*Peredibacter*	0.1	0±0 b	0±0 b	0±0 b	**0.3±0 a**
Firmicutes	***Bacilli***	***Bacillales***	*Bacillaceae***	*Bacillus***	0.3	**13.3±1 a**	11.5±1 b	8.3±4 c	12.2±3 ab
				*Lysinibacillus***	0	0.1±0 b	0.1±0 b	0.1±0 b	**5±2 a**
				Marinibacillus	0	**0.4±0 a**	0.1±0 b	0.1±0 b	0±0 b
				*Oceanobacillus***	0	**1.4±0 a**	**1.2±0 a**	0.7±0 b	0±0 c
			Paenibacillaceae	*Ammoniphilus*	0	**0.4±0 a**	0.2±0 a	0.3±0 a	0±0 b
				Cohnella	0	**0.3±0 a**	0.2±0 a	0.1±0 ab	0±0 b
				*Paenibacillus**	0.2	**1.1±0 a**	**1±0 ab**	0.8±1 ab	0.7±0 b
			*Planococcaceae***	*Paenisporosarcina*	0	**2.3±1 a**	1.4±0 b	1.9±1 a	0±0 c
				Sporosarcina	0	**1.5±0 a**	1.1±0 b	0.7±0 b	0±0 c
	*Clostridia**	*Clostridiales***	*Incertae Sedis XVIII*	*Symbiobacterium*	0	0.3±0 a	0.1±0 ab	0.2±0 a	0±0 b
			Ruminococcaceae	*Acetivibrio*	0	0.1±0 ab	0.2±0 a	0.1±0 ab	0±0 b
		Natranaerobiales	*Natranaerobiaceae*	*Dethiobacter*	0	0.2±0 a	0.2±0 a	0.1±0 ab	0±0 b
	Erysipelotrichi	*Erysipelotrichales*	*Erysipelotrichaceae*	*Turicibacter*	0	0.3±0 a	0.4±0 a	0.3±0 a	0±0 b
Actinobacteria	***Actinobacteria***	***Actinomycetales***	***Micrococcaceae***	*Arthrobacter*	0	**0.5±0 a**	**0.5±0 a**	0.1±0 b	0.3±0 ab
			Promicromonosporaceae	*Promicromonospora*	0.1	0±0 ab	0±0 ab	0±0 b	0.1±0 a
			Streptomycetaceae	*Streptomyces*	0.3	0.2±0 b	0.1±0 b	0±0 c	**1.8±0 a**
*Bacteroidetes**	*Sphingobacteria**	*Sphingobacteriales**	*Chitinophagaceae***	*Chitinophaga*	0	0±0 b	0±0 b	0.2±0 b	**0.6±1 a**
			*Cytophagaceae**	*Adhaeribacter**	0	**2.5±1 a**	**2.2±0 a**	1±1 b	0.5±0 c
				Pontibacter	0	2.2±1 c	3.8±2 b	**4.8±3 a**	0.5±0 d
			Sphingobacteriaceae	*Sphingobacterium*	0	0±0 b	0±0 b	0±0 b	0.4±0 a
Gemmatimonadetes	***Gemmatimonadetes***	***Gemmatimonadales***	***Gemmatimonadaceae***	***Gemmatimonas***	6.3	0.5±1 b	0.3±0 b	0.3±0 b	1.2±0 a

Note : Taxonomic group with significantly different community structure: *:0.01<p<0.05; **:0.001<p<0.01; bold taxa: p<0.001; bold number: soil with highest relative abundance of the taxa relative abundance=100x(number of sequences belonging to a taxa)/total number of sequences detected for the sample. Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; C: charcoal.

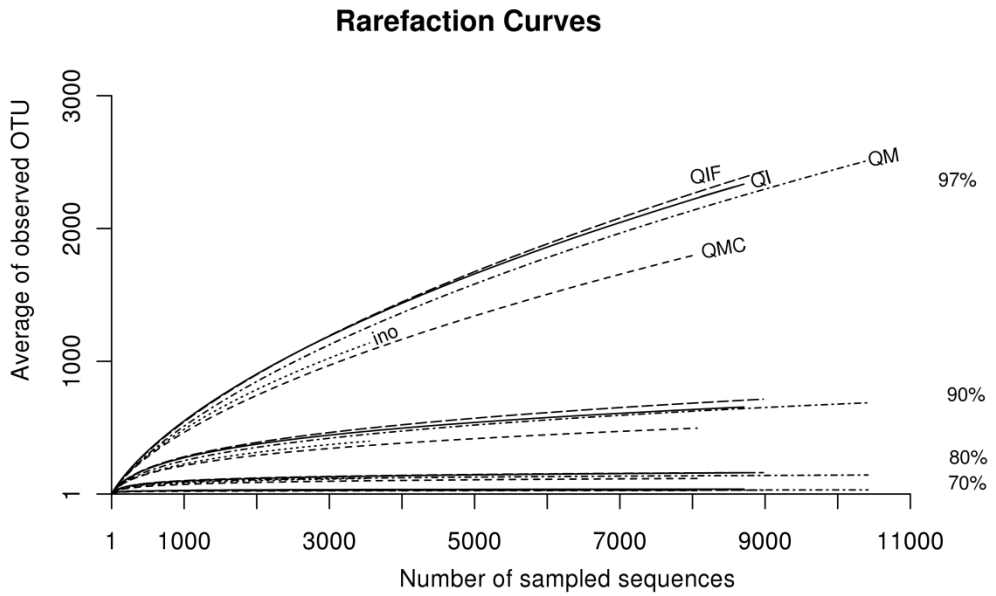

Fig. 1 Plot of rarefaction curves at different similarities. Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; C: charcoal.

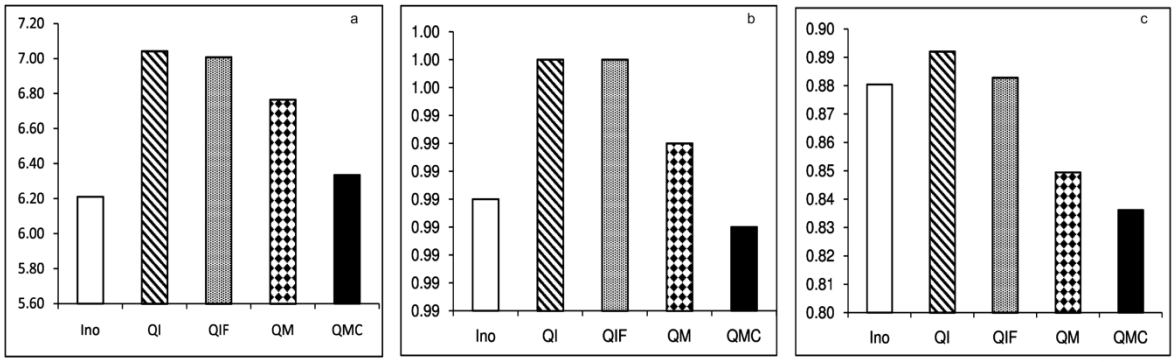

Fig. 2 Shannon (a), Simpson (b) diversity index and Pielou's evenness based OTUs (>97%) for different artificial soil samples. Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; C: charcoal.

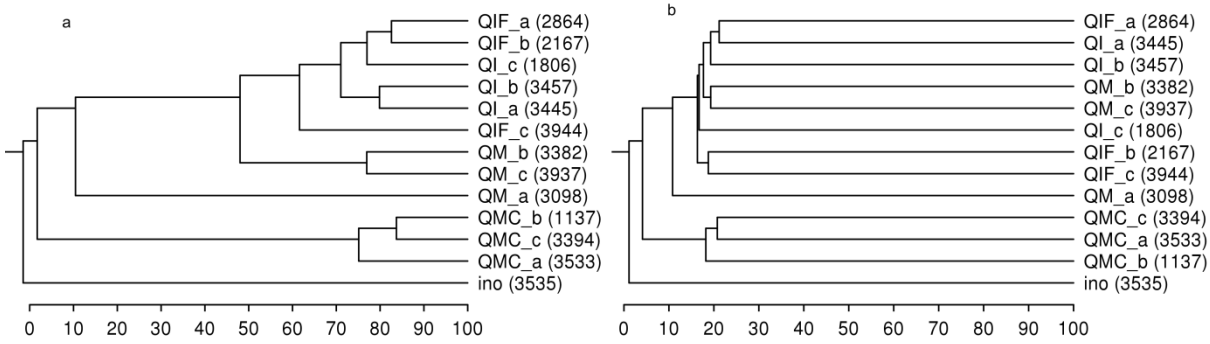

Fig. 3 Furthest-neigbor-joining cluster of bacterial community structure based on Pearson correlation distance (a) and Jaccard distance (b) using OTU (> 97%) as the species data. Number in brackets: number of sequences detected for the sample. Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; C: charcoal.

Table S1: Percent of different taxa detected for different artificial soil samples 90 days after incubation

		Ino	QI	QIF	QM	QMC
Proteobacteria	*Alphaproteobacteria*	42.4	9.7±1	8.5±2	8.7±3	18.5±1
	Betaproteobacteria	5.8	9.6±3	6.6±2	6±2	7.1±1
	Gammaproteobacteria	5.3	9.1±1	14.1±4	31.1±31	22.4±5
	Deltaproteobacteria	1.1	0.9±0	0.7±0	0.6±0	0.7±0
	Unclassified	0.9	0.1±0	0.4±0	0.3±0	1.4±0
Firmicutes	*Bacilli*	1.1	44.3±4	39.3±3	31.2±16	29.3±6
	Clostridia	0.1	5.6±1	6.7±0	4.2±2	0.1±0
	Erysipelotrichi	0	0.3±0	0.4±0	0.3±0	0±0
	Unclassified	0.1	0.8±0	0.7±0	0.5±0	0.1±0
Bacteroidetes	*Sphingobacteria*	1.1	9±2	11.1±2	9.9±5	6.9±1
	Flavobacteria	0	0.2±0	0.5±0	0.5±0	0.2±0
	Unclassified	0.3	0.7±0	1.5±1	0.8±0	0.9±1
Actinobacteria	*Actinobacteria*	19.2	5.1±1	5.1±1	3±1	6.7±1
Gemmatimonadetes	Gemmatimonadetes	6.3	0.5±1	0.3±0	0.3±0	1.2±0
Unclassified		9.2	3.6±1	3.6±0	2.3±1	3.9±3

Note: average±standard deviation; Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; C: charcoal.

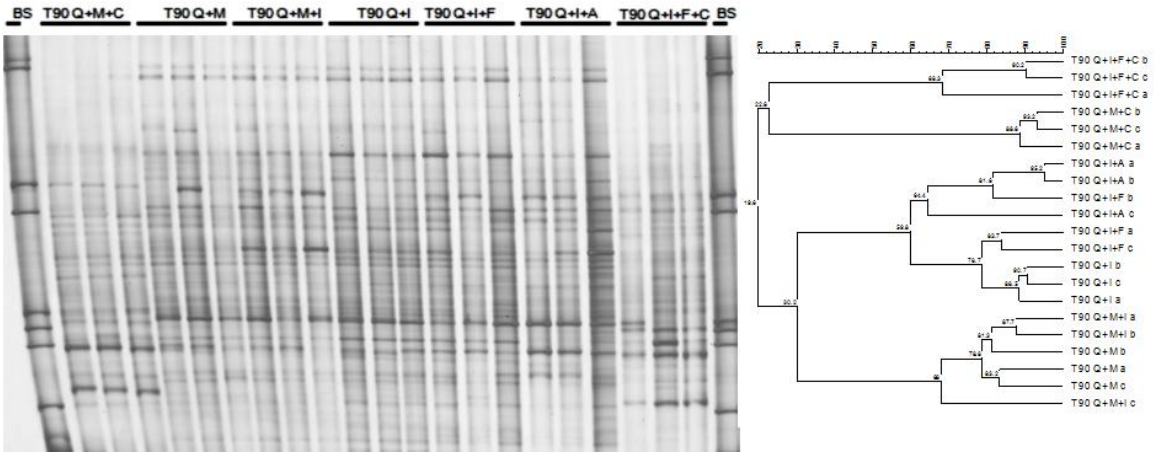

Fig. S1 Bacterial DGGE fingerprints and the corresponding UPGMA clusters for artificial soils at Day 90. Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; A: boehmite; C: charcoal.

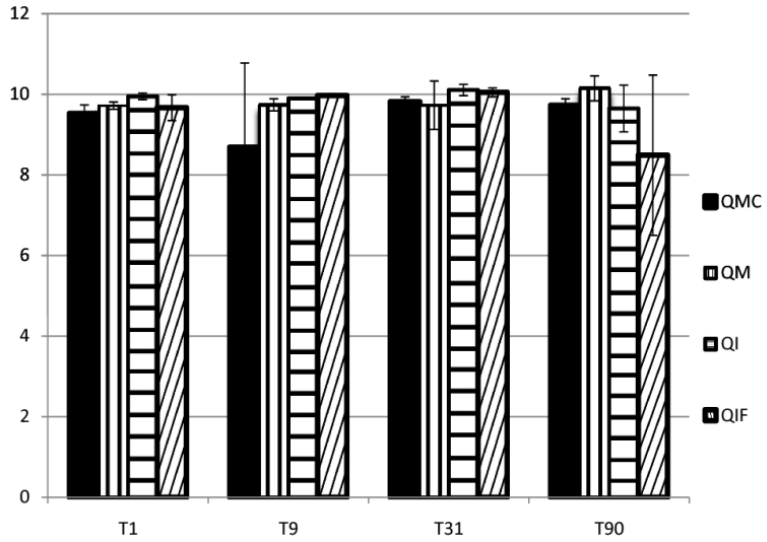

Fig. S2 Determination of 16S rRNA gene copy numbers (Log10 of 16S rRNA gene copy numbers per gram soil) for four artificial soil samples collected at Day 1, 9, 31 and 90. Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; C: charcoal.

Table S2: soil properties

soil composition	day 1	day 90		
	pH	OC mg g^{-1}	N mg g^{-1}	C/N
QM	7.7	14.6 ± 0.7	1.2 ± 0.1	11.7
QI	7.6	15.6 ± 0.6	1.3 ± 0.1	11.8
QMI	7.7	14.1 ± 0.9	1.2 ± 0.1	11.6
QMC	7.6	29.8 ± 2.0	1.4 ± 0.2	20.7
QIF	7.6	17.8 ± 3.8	1.6 ± 0.3	11.4
QIA	7.6	14.7 ± 0.4	1.3 ± 0.1	11.6
QIFC	7.6	31.2 ± 2.7	1.6 ± 0.3	19.2

Note: Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; A: boehmite; C: charcoal.

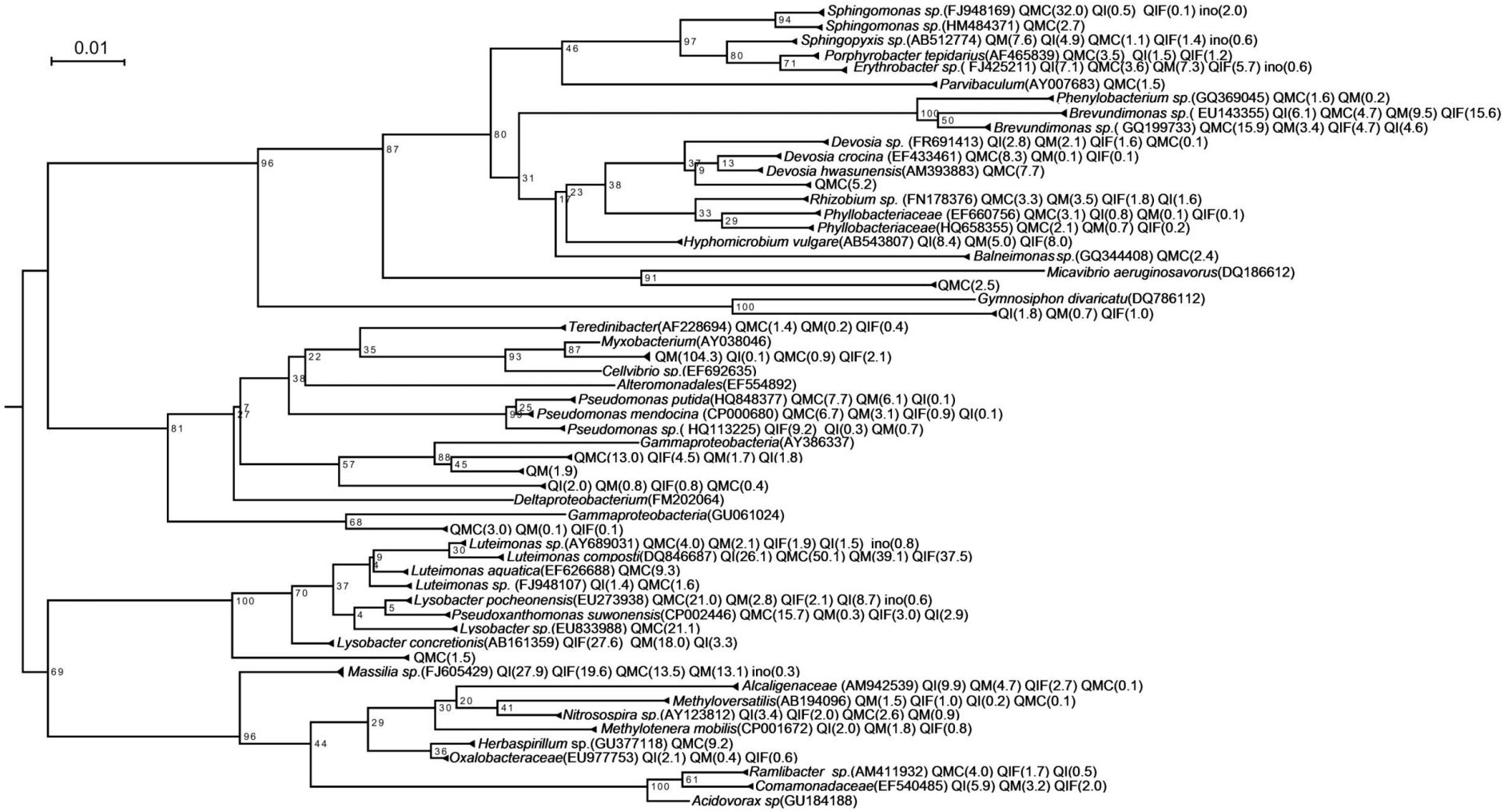

Fig. S3: Neighbor-joining rooted phylogenetic tree based on multiple alignments of representative sequences for discriminative OTUs belonging to *Preteobacteria*. Values at each node = (bootstrap value/100) x 500. Number in brackets: Relative abundance in ‰.

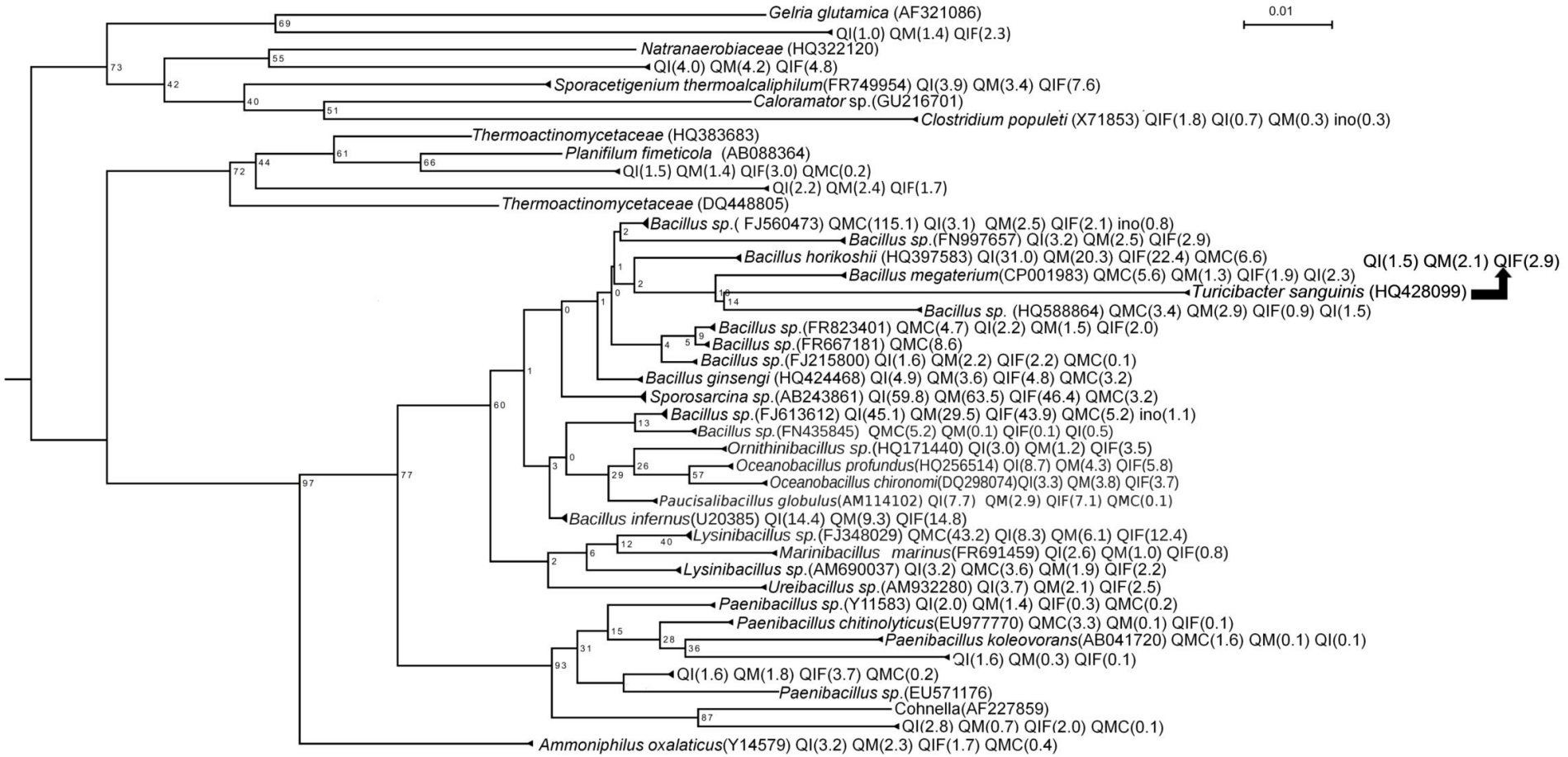

Fig. S4: Neighbor-joining rooted phylogenetic tree based on multiple alignments of representative sequences for discriminative OTUs belonging to *Firmucutes*. Values at each node = (bootstrap value/100) x 500. Number in brackets: Relative abundance in ‰.

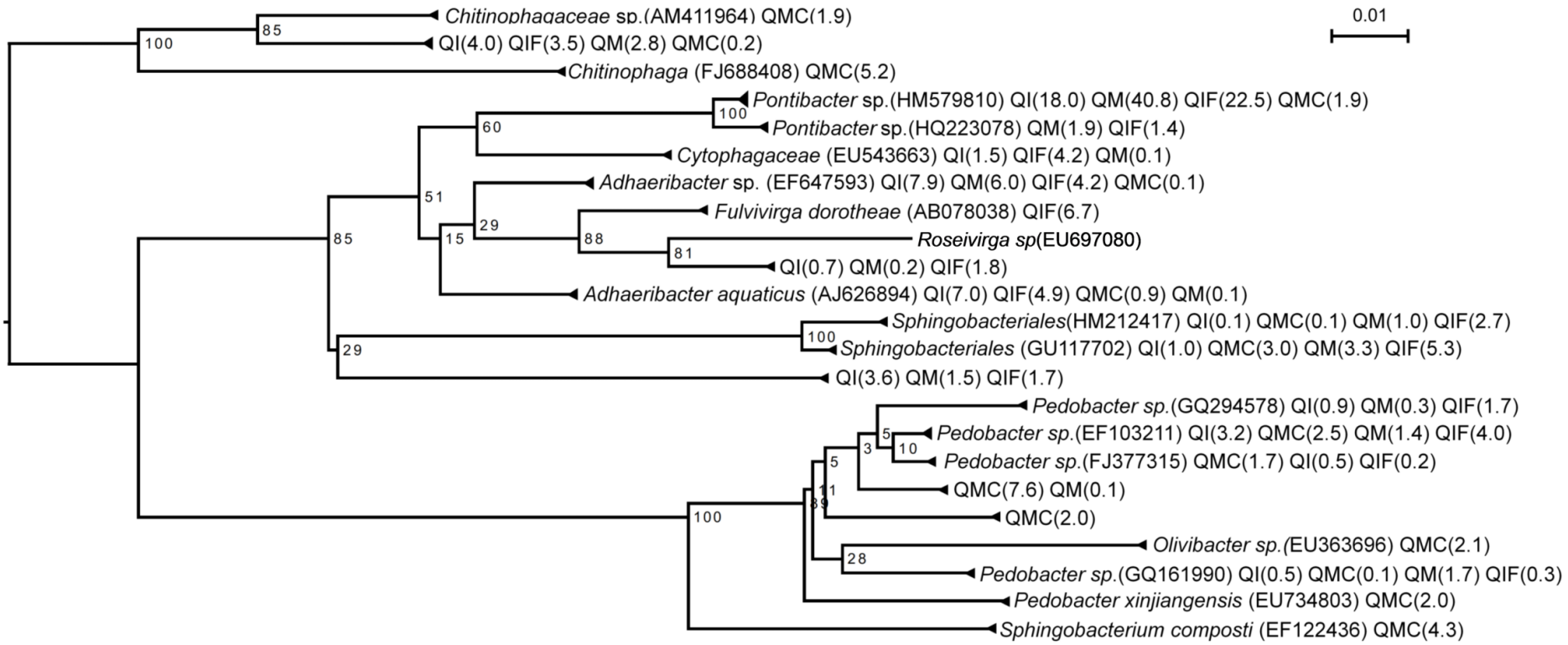

Fig. S5: Neighbor-joining rooted phylogenetic tree based on multiple alignments of representative sequences for discriminative OTUs belonging to *Bacteroidetes*. Values at each node = (bootstrap value/100) x 500. Number in brackets: Relative abundance in ‰.

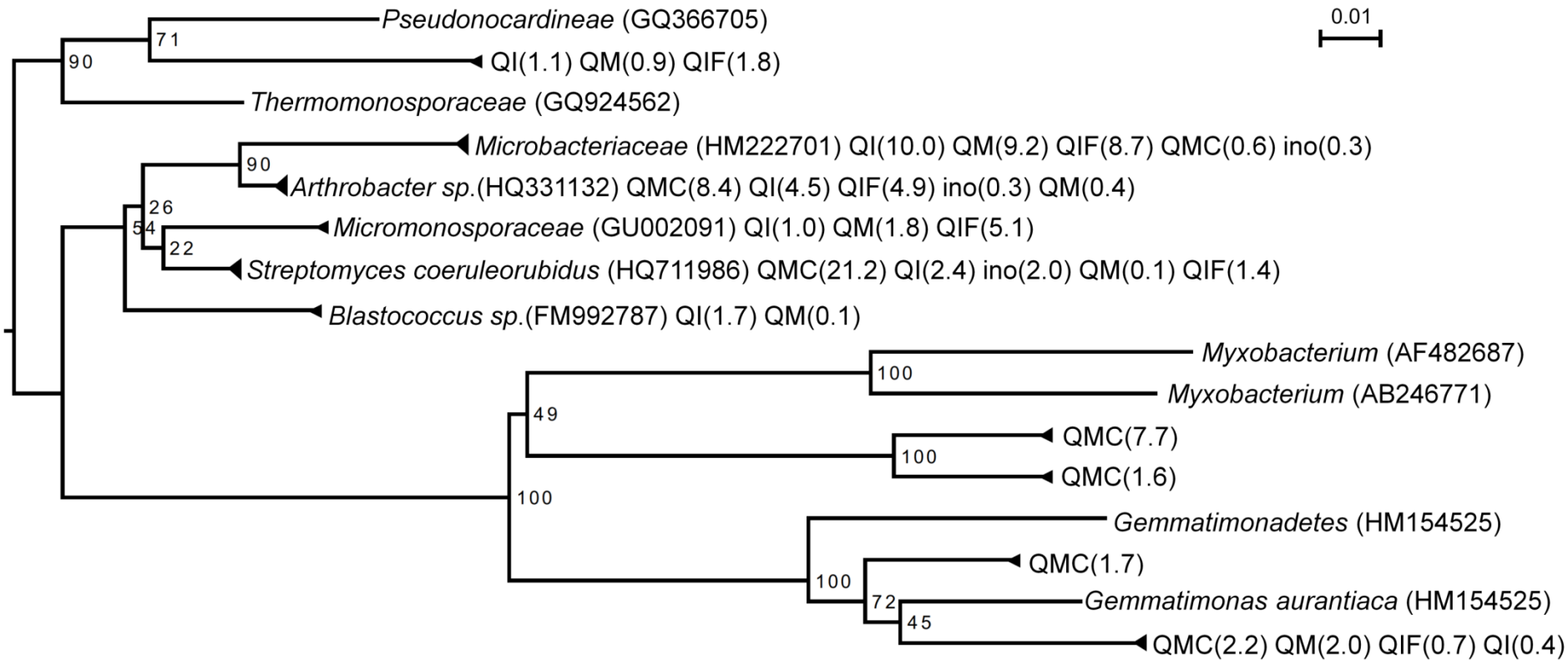

Fig. S6: Neighbor-joining rooted phylogenetic tree based on multiple alignments of representative sequences for discriminative OTUs belonging to *Actinobacteria* and unidentified phylum. Values at each node = (bootstrap value/100) x 500. Number in brackets: Relative abundance in ‰.

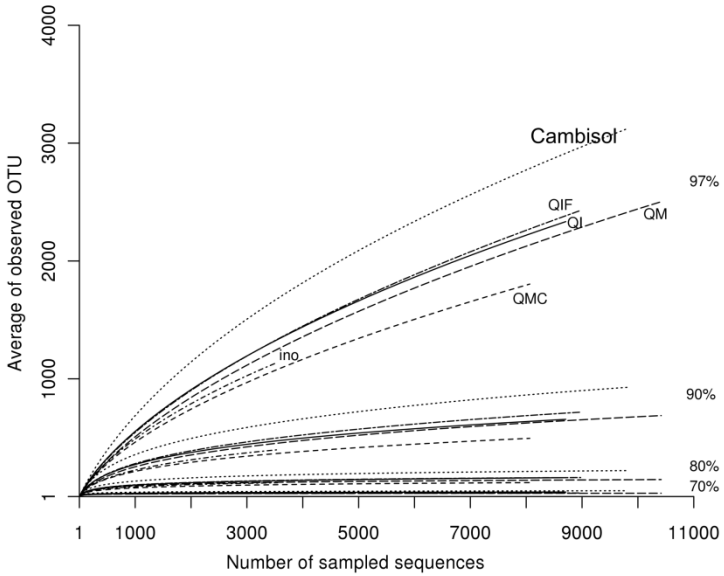

Fig. S7: Plot of rarefactions curves for cambisol, QIF, QI, QM, ino and QMC at different levels. Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; A: boehmite; C: charcoal.

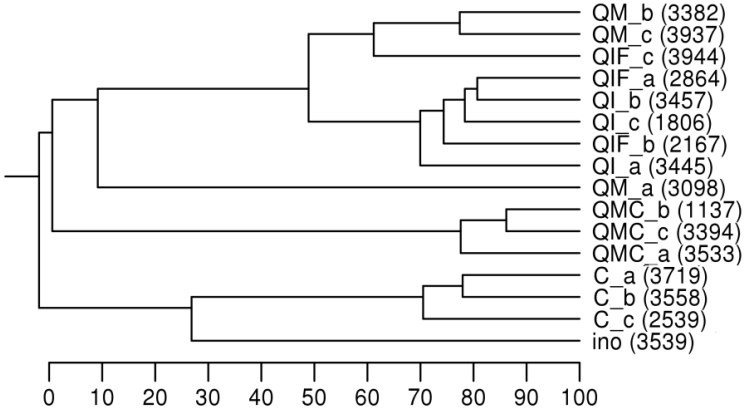

Fig. S8. Cluster analysis of bacterial community structure based on Pearson correlation index distance. Ino: inoculant, Q: quartz; I: Illite; M: montmorillonite; F: ferrihydrite; A: boehmite; C: charcoal.

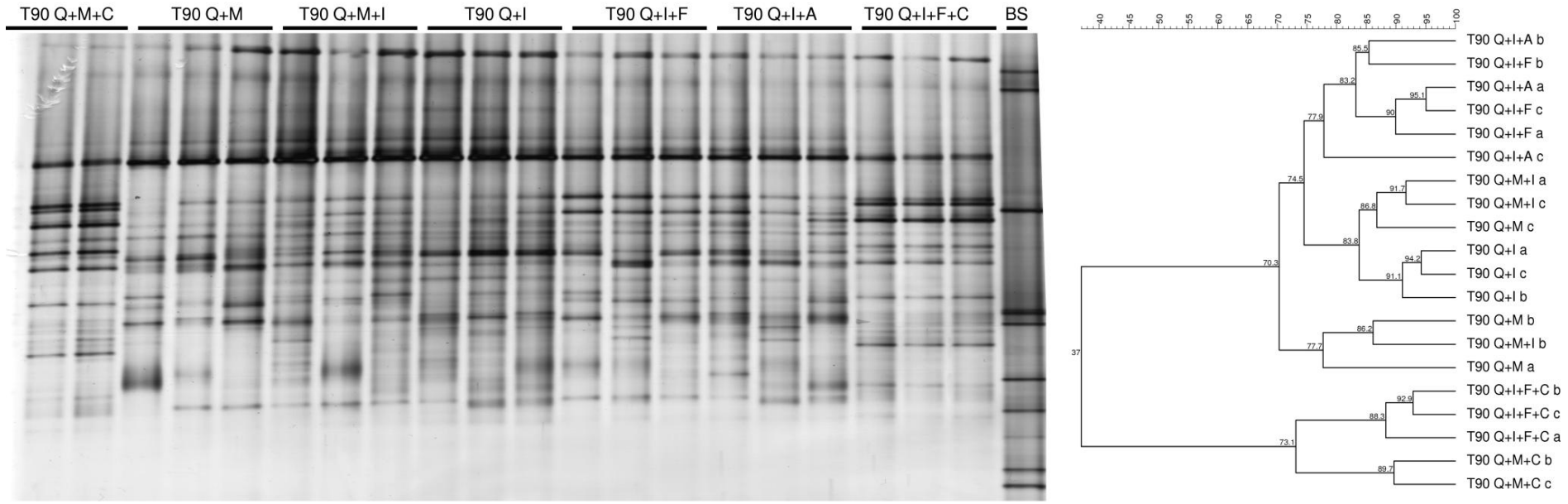

Fig. S9 *Betaproteobacteria*-specific DGGE fingerprints and the corresponding UPGMA clusters for different artificial soils spiked at day 90.

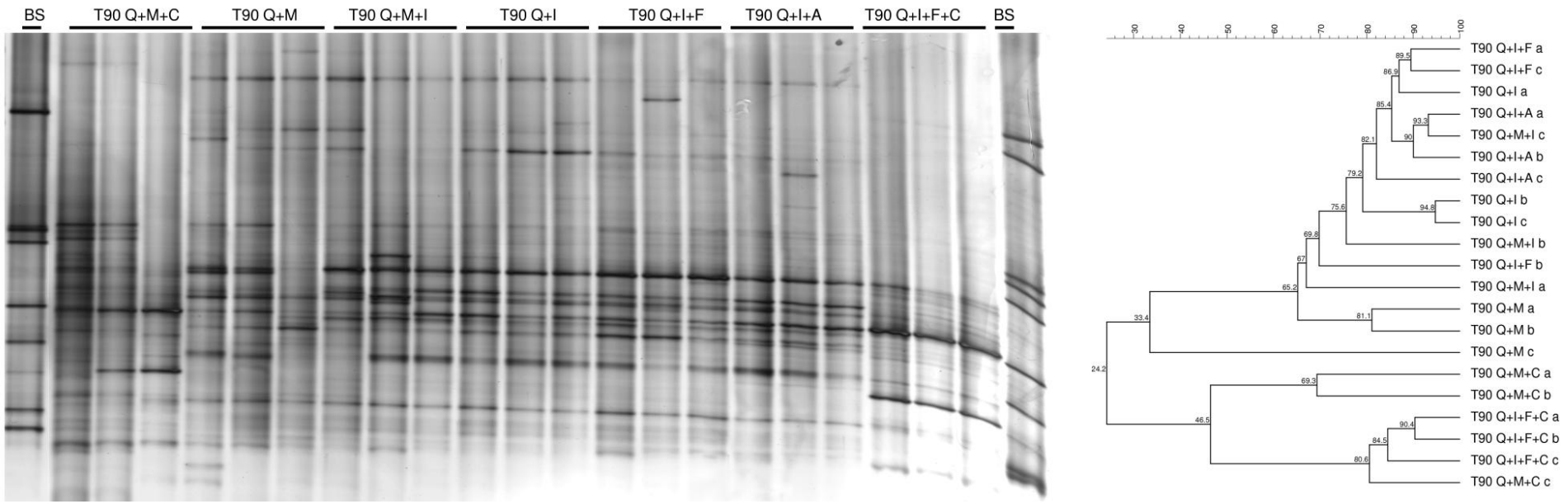

Fig. S10 *Actinobacteria*-specific DGGE fingerprints and the corresponding UPGMA clusters for different artificial soils spiked at day 90.

Outlook

So far, the responses to phenanthrene in two soils of bacteria harboring *PAH-RHDα* genes, indigenous microbial communities, functional microbial groups and IncP 9 plasmid were explored. Molecular tools and data manager pipelines for geochip, phylochip and 16S rRNA gene amplicon pyrosequencing were constructed. Also, the roles of minerals on the assembling of microbial communities were evaluated.

To study the characterization of biogeochemical interfaces, an integrated microcosm experiment is planned or presently running. Soil physico-chemical characters as well as microbial community structure will be measured. Huge amounts of interdiscipline data concerning physical, chemical and biological processes are expected. To interlink between these processes, not only great mathematical skills but also close cooperation between soil physician, soil chemist and microbial ecologist are required to properly understand the power and limitations of the tools applied.

In addition, techniques to precisely quantify bacteria or certain taxa based on minor amounts of samples could be of interest. From the biological aspect, precisely parallel quantification of several genes or functions is looked forward to. It was also noticed that the fast developing high throughput sequencing technology could provide more insight into the microbial composition, functional potential or biochemical pathway. Basic ecological questions, such as “Who are they? What are they doing? Who does what?” will be more properly answered. But of course, novel tools (hardware and software) are needed to enjoy the data banquet.

In summary, it is really a great luck to witness the fast evolution of microbial ecology as always new knowledge, new conceptions and new tools appears. This fantastic development provides a lot of fun motivating and cheering us to think, to change, to search and to research differently.

Acknowledgements

Now, it is time to say thanks to many people who supported, guided, encouraged and contributed to this work directly and indirectly.

First of all, I would like to thank the Deutsche Forschungsgemeinschaft (DFG) for financially supporting this project which is the prerequisuit for every nice thing following. Equally, I am quite grateful for my supervisor. Professor Dr Kornelia Smalla for her excellent guidance, several excellent chances, great patience, sharp criticism and impressive mechanisms for quality controlling, constructive work on the papers and manuscripts.

Secondly, I thank Dr Holger Heuer thousand times because of his sharing of sharp and deep views on statistics, primer design and data management, constructive suggestion and comments on the thesis as well as your nice cool jokes.

A lot of thanks to all members who belong to or belonged to the group of AG Smalla or the colleagues in the institute of Epidemiology and Pathogen Diagnostics specially for Mrs. Jungkurth who continually made wonderful corrections on my manuscripts, abstracts and posters; also to Ellen Krögerrecklenfort who initiatlly reformed my quality standards for practical lab work and to Newton C.M.Gomes who shared his important experiences on DGGE. Also to Alexandra Moura, Christoph Kopmann, Flavia Dematheis, Susanne Schreiter, Simone Dealtry-Gomes, Maria Touceda, Viola Weichelt, Sven Jechalke, Ute Zimmling, Nicole Weinert, Sarah Xenia Meyer, Kerstin Dowideit, Doreen Babin et al. for their contributions to a warm and peaceful atmosphere during and after working.

A lot of thanks also to all members in the giant SPP1315 project especially to our close cooperators Katja Heister, Geertje Johanna Pronk, Ingrid Kögel-Knabner, Sebastian Zühlke and Michael Spiteller for their excellent work. The meetings were always very enjoyable and inspiring. Here, I would like to especially thank Professor Siegfried Kropf for his guidance and helping in statistics and Professor Jiuzhong Zhou for his generosity accepting me to work in his group with the Geochip.

Thanks also to all friends in China for their encouragment, helping, caring during special period.

Finally, I would like to give my deepest thanks to my family especially my parents for all their possible support during my education.

Curriculum Vitae

Guo-chun Ding: Jiangsu, China; 20th March, 1981. guochunding@gmail.com

Education

2007/10---present: PhD student at the Julius Kühn-Institute Braunschweig and at the Technical University Braunschweig, Germany under the supervision of Prof.Dr Kornelia Smalla.

2003/9---2007/6: Master student at Nanjing Agricultural University, China under the supervision of Prof. Jianhua Guo.

List of publications

Published

1. Ding GC, Heuer H, Zuhlke S, Spiteller M, Pronk GJ, Kogel-Knabner I and Smalla K. (2010) Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes by using a novel PCR detection system. Appl Environ Microbiol 76: 4765-4771.

2. Weinert N, Piceno Y, Ding GC, Meincke R, Heuer H, Berg G, Schloter M, Andersen G and Smalla K. (2011) PhyloChip hybridization uncovered an enormous bacterial diversity in the rhizosphere of different potato cultivars: many common and few cultivar-dependent taxa. FEMS Microbiol Ecol 75: 497-506.

In preparation or submitted:

1. Ding GC, Xie JP, J-Z Zhou, K Smalla. Responses of microbes in two typical central European soils to phenanthrene spiking as revealed by GeoChip (submitted)
2. Ding GC, Heuer H and Smalla K. Similar ribotypes belonging to *Sphingomonadales* and *Burkholderiales* enriched in two different phenanthrene spiked soils as revealed by pyrosequencing and DGGE (submitted)
3. Ding GC, Pronk GJ, Heuer H, Heister K, Kögel-Knabner I and Smalla K. Microbial community assembled on artificial soils as studied by PCR-DGGE and pyrosequencing

4. Ding GC, Heuer H and Smalla K. Establishment of a novel primer system targeting conserved stretches in the *rep-oriV* region of IncP-9 plasmids

Supplements

Bacterial diversity in the rhizosphere of potato plants: site- and cultivar-dependent responders revealed by PhyloChip analysis

Nicole Weinert[1a], Yvette Piceno[2a], Guochun Ding[1a], Remo Meincke[3], Holger Heuer[1], Gabriele Berg[3], Michael Schloter[4], Gary Andersen[2] and Kornelia Smalla[1]*

[1]Julius Kühn-Institut - Federal Research Centre for Cultivated Plants (JKI), Institute for Epidemiology and Pathogen Diagnostics, Messeweg 11-12, 38104 Braunschweig, Germany
[2]Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

[3]Graz University of Technology, Environmental Biotechnology, Petersgasse 12, A-8010 Graz, Austria
[4] Helmholtz Zentrum München, Department for Terrestrial Ecogenetics, Ingolstädter Landstr. 1, 85758 Oberschleissheim, Germany

Key words: 16S rRNA genes/community DNA/ clone library/ PhyloChip/potato rhizosphere; ,

*Correspondence

Kornelia Smalla, Julius Kühn-Institut, Institut für Epidemiologie und Pathogendiagnostik, Messeweg 11-12, D-38104 Braunschweig

[a]These authors equally contributed to this work**.**

Abstract

The taxonomic composition of bacterial communities in the rhizosphere of three different potato cultivars grown at two field sites was determined based on the analysis of 16S rRNA gene fragments amplified from total community DNA by employing high-density PhyloChips and sequencing six clone libraries. A total of 2432 operational taxonomic units (OTUs) were detected in the rhizosphere by PhyloChip hybridization. OTUs responding to the site (692) and the cultivar (207) or both (109) were identified. More than 40% of OTUs belonging to *Bradyrhizobiales, Sphingomonadales, Burkholderiales, Rhodocyclales, Xanthomonadales* and *Actinomycetales* significantly differed in their abundance between the sites. A relatively high proportion of OTUs belonging to the *Pseudomonadales, Actinomycetales* and *Enterobacteriales,* were identified as taxa responding to the cultivar. Highly significant effects of the site were detected for all major phyla while only the *Actinobacteria* and *Firmicutes* revealed significant differences in response to the cultivar. The finding that the abundance of *Pseudomonadales* and *Actinomycetales* was significantly affected by the cultivar is remarkable in view of the importance of strains belonging to these taxa as biocontrol strains. The analysis of 16S rRNA gene clone libraries showed that the majority of ribotypes detected in the rhizosphere of flowering potato plants from both sites were affiliated to the phylum *Proteobacteria* with *Alphaproteobacteria* representing the most frequently detected class. Furthermore, the 16S rRNA gene sequences of the dominant clones of all cultivars from both sites were affiliated to the same phyla and classes and that the majority had the highest sequence identity to uncultured bacteria.

Introduction

Microbial communities colonizing the rhizosphere of crop plants are of central importance not only for plant nutrition, health and quality but also for carbon sequestration and nutrient cycling in terrestrial ecosystems. Biotic and abiotic factors are assumed to shape the structural and functional diversity of microbial communities in the rhizosphere (Berg and Smalla, 2009). Soil properties, plant species, the cultivar and plant growth stages were shown in several studies by means of cultivation-dependent and -independent methods to influence the composition of rhizosphere microbial communities (Lemanceau *et al.,* 1995; Marschner *et al.,* 2001; Smalla *et al.,* 2001; Berg et al., 2002, 2006; Briones *et al.*, 2002; Heuer *et al.,* 2002; Kowalchuk *et al.,* 2002; Costa et al., 2006; 2007; Bremer et al., 2007; Weinert et al., 2009a,b). Cultivation-dependent studies often focused on bacterial isolates with *in vitro* antagonistic activity towards fungal pathogens. The molecular characterization of antagonists and their identification revealed both site- and plant-dependent diversity (Berg *et al.,* 2002; 2006). Frequently obtained antagonists isolated from different sites often belonged to the genera *Pseudomonas*, *Streptomyces*, *Bacillus* or *Serratia*. As soil bacteria that easily form colonies on solid nutrient media represent only a minor fraction of the rhizobacteria, the analysis of DNA extracted directly from soil adhering to the plant roots or from the microbial pellet obtained from roots became more and more popular over the last decade. To gain insights into abundant ribotypes or ribotypes of particular taxa, 16S rRNA genes amplified from rhizosphere total community (TC) DNA are analyzed by molecular fingerprints or by cloning and sequencing. Molecular fingerprints such as automated ribosomal intergenic spacer analysis (ARISA), denaturing gradient gel electrophoresis (DGGE), single strand conformation polymorphism (SSCP) or terminal restriction analysis (T-RFLP) have the great advantage that a rapid comparative analysis of multiple samples can be performed which is essential for testing treatment effects and elucidating the temporal dynamics of bacterial communities in response to environmental gradients or experimental factors (Muyzer and Smalla, 1998; Fisher and Triplett, 1999; Forney *et al.,* 2004; Kowalchuk *et al.,* 2006; Smalla *et al.*, 2007). While the analysis of 16S rRNA gene fragments amplified from rhizosphere TC DNA by means of these fingerprinting techniques is an ideal tool to follow succession or effects of soil types or plants on the microbial community composition, taxonomic information on the major players and how they respond to these factors are difficult to obtain. Although

ribotypes responding to treatment effects can be identified after cloning and sequencing of re-amplified bands (Heuer *et al.*, 2002; Costa *et al.*, 2006), molecular fingerprints are certainly not the appropriate approaches to gain deeper insights into the bacterial diversity in the rhizosphere. Information on the taxonomic composition of rhizosphere communities can be gained by sequence analysis of cloned 16S rRNA genes amplified from TC DNA. However, costs often prohibit the analysis of sufficient numbers of replicates and numbers of clones. To increase the throughput of the detection of bacteria in complex samples such as soil, high density 16S rRNA gene probe arrays have been developed. The so-called PhyloChips developed by Brodie *et al.* (2007) allow the comprehensive detection and comparison of *Bacteria* in complex environmental samples. Recently, the applicability of the PhyloChip was also explored for soils and oat rhizosphere (Cruz-Martínez *et al.*, 2009, DeAngelis *et al.*, 2009). The PhyloChips used offer the potential to detect 8,741 known OTUs simultaneously. Here we employed for the first time high-density PhyloChips to explore the bacterial diversity in the rhizosphere of three different potato cultivars grown at two sites with different soil characteristics. 16S rRNA gene fragments were amplified from the total community DNA of three replicates per cultivar and site. The amplicons were subsequently fragmented, biotinylated and hybridized with the PhyloChip. In parallel, one clone library was generated per cultivar and site from the pooled 16S rRNA gene amplicons. While cloning and sequencing of 16S rRNA gene fragments was done to identify abundant ribotypes in the rhizosphere of flowering potato plants, statistical analysis of the PhyloChip data enabled us to determine site- and cultivar-dependent responders. MANOVA analysis of the first five principal components at different taxonomic levels (phylum, class, order to family) were used to test the PhyloChip data for significant effects of the site and the cultivar. Major taxa contributing to the differences between the sites and among the cultivars were identified.

Materials and methods

Field design and sampling

In 2006, three commercial potato (*Solanum tuberosum* L.) cultivars, 'Baltica', 'Désirée' and 'Sibu' were grown in a randomized field trial at two sites in Southern Germany, Roggenstein and Oberviehhausen. Each cultivar was grown in six replicated plots. Samples taken at flowering stage (EC60 according to Hack *et al*., 1993) were analyzed. Each replicate composite sample consisted of the roots of five plants taken from one plot. The potato plants were carefully removed and shaken to detach loosely adhering soil. Roots were cut into pieces, combined and immediately brought to the laboratory. For each cultivar, four composite samples were processed. Further details on the soil properties and experimental design are given by Weinert *et al.* (2009a).

Sample processing and extraction of total community (TC) DNA

The microbial cells associated with the roots were obtained from 10 g of root pieces which were placed in a sterile Stomacher bag. After addition of 30 ml Milli-Q water the samples were subjected to a homogenization step for 60 s at high speed using a Stomacher laboratory blender (Seward, West Sussex, UK). This homogenization step was repeated three times, and microbial cells were harvested from the resulting suspensions by centrifugation at 10,000 x *g* for 15 min at 4 °C. The microbial pellets were frozen at -80 °C until TC DNA extraction. TC DNA was extracted from 0.5 g of the cell pellet after a harsh lysis step (FastPrep FP120 bead beating system Q-Biogene, Carlsbad, CA, USA) by means of the BIO-101 DNA spin kit for soil (Q-Biogene). The DNA was purified using the Geneclean Spin Kit (Q-Biogene) according to the manufacturer's instructions, except that the DNA was eluted in the same volume that was used for purification. TC DNA was 1:5 diluted with elution buffer of the purification kit and stored at -20 °C until further processing.

Cloning and sequencing of 16S rRNA gene fragments

The TC DNA of four replicates per cultivar were used to amplify approximately 1450 bp of the 16S rRNA gene using the primer pair F27 (5'-AGAGTTTGATC(A/C)TGGCTCAG -3') and R1494 (5'-CTACGG(T/C)TACCTTGTTACGAC -3') (Weisburg *et al*., 1991; Heuer *et al.*, 2009). PCR reaction mixtures (25 µl) consisted of 1 µl of template DNA (1-5ng), 1 x Stoffel

buffer II (Applied Biosystems), 0.2 mM dNTPs, 3.75 mM $MgCl_2$, 2 % (v/v) dimethylsulfoxide (DMSO), 2.5 µg bovine serum albumin (BSA), 0.1 mM of each primer and 2.5 U Taq DNA polymerase Gold (Applied Biosystems). PCR-conditions consisted of an initial denaturation step at 94 °C for 5 min, followed by 25 cycles of 94 °C for 1 min, 56 °C for 1 min and 72 °C for 2 min and a final extension step at 72 °C for 10 min. Five µl of the PCR-products were checked on a 1 % agarose gel in 0.5 x TBE buffer at 80 V, stained with ethidium bromide and photographed under UV. A total of 60 µl of PCR products were obtained for each cultivar by combining 15 µl per replicate. The PCR products were purified using the Geneclean Spin Kit (Q-Biogene) according to the manufacturer's instructions. After purification the yield was checked on an agarose gel and 3 µl of purified PCR-product were ligated into pGEM-T vectors (Promega, Madison, USA) and transformed into competent cells (*Escherichia coli* JM109, Promega) as described by the manufacturer. After overnight incubation at 28 °C, 100 white colonies per cultivar and site were picked with a sterile tooth pick and streaked out on fresh medium. The colonies were then transferred into sterile Milli-Q water, boiled at 100 °C for 10 min and afterwards stored on ice. The primer pair SP6/T7 was used to screen for clones with correct insert length. PCR-products of the expected length were purified with the Geneclean Spin Kit and sent for sequencing (IIT Biotech GmbH, Bielefeld, Germany).

Sequence analysis

Sequence analysis of the clones was performed using both the seqmatch analysis tool of the Ribosomal Database Project II, release 9.61 (http://rdp.cme.msu.edu/) and the nucleotide-nucleotide BLAST search tool (BLASTN) of the National Center for Biological Information (NCBI, USA). Multiple alignment of DNA sequences was done with clustalx2(2.0) and the DNA distance was calculated with software package phylip(3.68). The program DOTUR (distance-based operational taxonomic unit [OTU] and richness determination) was used for rarefaction analysis.

PCR amplification of 16S rRNA genes for PhyloChip hybridization

TC DNA extracts of three replicates per cultivar and site were amplified using an 8-temperature gradient PCR and universal 16S rRNA primers
27f (5'- AGAGTTTGATCCTGGCTCAG -3) and 1492r (5'- GGTTACCTTGTTACGACTT -3'). One µl of template (1-5 ng) was used at each

temperature. 25 µl reactions (final concentration was 1x Ex Taq Buffer with 2 mM $MgCl_2$, 300nM each primer (27f and 1492r), 200 µM each dNTP (TaKaRa), 25 µg bovine serum albumin (Roche Applied Science, Indianapolis, IN), and 0.625 U Ex Taq (TaKaRa Bio, Inc., through Fisher Scientific, Pittsburg, PA)) were amplified using an iCycler (Bio-Rad, Hercules, CA). Eight replicate PCR amplifications were performed at a range of annealing temperatures from 48 to 58 °C with an initial denaturation at 95 °C for 3 min followed by 25 cycles of 95 °C (30 s), annealing (30 s), and 72 °C for 2 min, and then final extension for 10 min at 72 °C. PCR products from each annealing temperature of a sample were combined and concentrated to 40 µl or less final volume using Microcon YM-100 filters (Millipore, Billerica, MA). 40 µl water was added to the filter units and spun through prior to loading the PCR product for concentration. One µl of concentrated PCR product was quantified on a 2% agarose E-gel using the Low Range Quantitative DNA Ladder (Invitrogen, Carlsbad, CA).

500 ng PCR product was applied to each PhyloChip (G2) following the previously described procedures (DeSantis *et al.*, 2007) with the following change: the biotinylation was performed using GenChip Labeling Reagent (Affymetrix) and terminal deoxynucleotidyl transferase (Promega) according to the manufacturer's (Affymetrix) recommendations.

Statistical analysis

OTU-level report was produced mainly as Brodie *et al.* (2007) described for background subtraction, detection and quantification criteria.The exception is the normalization of array data which was normalized according to the average total array intensity. The signal intensities of those not detected OTUs were shifted to 1 to avoid errors in subsequent log10 transformation. Software package R (2.8.0) was used for statistical analysis and summary reports. Based on signal intensities, two-way ANOVA (analysis of variance) was used to test the signal intensity of each OTU in order to identify OTUs which significantly (unadjusted $p < 0.05$) differ in response to cultivar and/or sites. Numbers of detected OTUs, as well as those significantly responding to treatments, were summarized at different taxonomic levels. To find the discriminating OTUs ($p < 0.05$) between cultivars, Tukey's test was applied in conjunction with two way ANOVA. The result was also summarized at different taxonomic levels. The treatment effect on high level taxonomic groups such as

phylum, class, order and family, were tested basically according to Glimm *et al*. (1997) the fluorescent signal intensity of each OTU was log10 transformed, centered (to reach a 0 mean) standardized (to reach a standard deviation of 1) and assigned to different taxonomic groups. These adjusted data are the basis for subsequent analysis. Principal Component Analysis (PCA) was performed with the PCA function within R package named “FactoMineR”. Matrix of sample coordinates on the first five components was tested with Wilks'Λ. In this case, the test is equivalent to F-test with degrees of freedom 5 and 8 for site and 10 and 16 for cultivar.

Nucleotide sequence accession numbers

The nucleotide sequences of cloned 16S rRNA gene fragments determined in this study were deposited in the GenBank database under accession numbers FJ845053-FJ845364.

Results

PhyloChip-based bacterial community analysis

PhyloChips comprising 350,000 oligonucleotides were employed to gain insights into the structural diversity of bacteria in the rhizosphere of flowering potato plants of three different cultivars grown at two field sites. Biotinylated 50-200 bp fragments obtained from 16S rRNA gene fragments amplified from TC DNA of three independent replicates per cultivar were hybridized with the PhyloChips. A total of 2432 OTUs gave a detectable hybridization signal on the PhyloChips (pf>=0.9). An overview of the taxa detected per each cultivar and site and the corresponding standard deviation is given in Table 1. The total number of OTUs detected per cultivar ranged from 1444±221.7 (Baltica, Oberviehhausen) to 2015±138.2 (Baltica, Roggenstein). In total 43 phyla were detected (Table S1). However, only 13 phyla with ten or more OTU were observed. The highest number of OTUs belonged to the *Proteobacteria* (1124) followed by *Firmicutes* (440), *Actinobacteria* (269), *Bacteroidetes* (176) and *Acidobacteria* (75). Among the proteobacterial OTU the highest numbers were affiliated to *Gamma*- (399) followed by *Alpha*- (351), *Beta*- (203), *Delta*-(115) and *Epsilonproteobacteria* (47). The total number of OTUs detected per order or family was often much higher than the number of OTUs detected per cultivar of each site indicating that some OTUs were not detected at

both sites. The numbers of OTUs that were detected in more than five of the nine replicates at one site and in less than four of the nine replicates at the other site were considered at site-specific OTUs and are listed in Table 2. A considerably higher number of so-called site-specific OTUs was detected in the rhizosphere of potato cultivars grown at the Roggenstein site (194) compared to the Oberviehhausen site (17). Site-specific OTUs for Roggenstein were affiliated to the *Actinobacteria* (19), the *Burkholderiales* (15), *Clostridiales* (15) and *Rhizobiales* (12). However, when only OTUs were considered that were detected in all three replicates of a treatment the total number of OTUs detected was 1,856. The relative proportion of the ten major phyla was similar for all three cultivars at both sites (Fig. 1).

Multiple two-way ANOVA analyses were applied to identify OTUs discriminative to site and cultivar, so-called responding OTUs. The numbers of OTUs that significantly differed depending on the site and the cultivar are summarized in Table 3. In total, 692 (28.5%) of the 2,432 OTUs significantly ($p<0.05$) differed between the sites, while only 207 (8.5%) OTUs significantly differed among the cultivars. Among these OTUs, 109 (4.5%) revealed cultivar-dependent differences at both sites.

Interestingly, specific responses to sites were observed for many different taxonomic groups. A high proportion (>40%) of detected OTUs belonging to the *Bradyrhizobiales, Sphingomonadales, Comamonadaceae, Burkholderiaceae, Rhodocyclaceae, Xanthomonadaceae, Microbacteriaceae, Streptomycetaceae, Micrococcaceae, Pseudonocardiaceae, Acidobacteriales* displayed a significant response to sites. A relatively high proportion of the OTU belonging to the *Enterobacteriales* (17.2 %), *Pseudomonadales* (39.6 %), *Micromonosporaceae* (52%), *Streptomycetaceae* (83.3) and *Bacilliales* (11.6%) responded to the cultivar. In order to identify the OTUs that significantly differ between cultivars more in detail, the Tukey´s honest test was performed for all OTUs and the results are given in Table 4. The signal intensities for 10 out of 19 detected OTUs belonging to the *Micromonosporaceae* were significantly higher for the cultivar Sibu compared to Désirée. Out of the 18 detected OTUs affiliated to the *Streptomycetaceae* the signal intensities of 13 OTUs were significantly higher for Sibu compared to Désirée and of 4 OTU for Sibu compared to Baltica. OTUs belonging to the *Baciliaceae* showed for 14 of the 91 detected OTU significantly higher signal intensities for Désirée compared to Baltica. The abundance of 19 and 15 OTU belonging to *Pseudomonadales* was

significantly higher in the rhizosphere of Désirée in comparison to Baltica and Sibu, respectively. Also the signal intensities of 7 out of 53 detected OTUs belonging to the *Enterobacteriales* were significantly higher for Désirée in comparison to Baltica and Sibu. In addition, the responses of different taxonomic groups at high levels (phylum, class, order and family) to the site and the cultivar were tested by means of MANOVA analyses of first five principal components. For all taxonomic levels (listed in Table 1) significant effects of the site were found except for *Clostridia*, *Enterobacteriales* and *Micromonosporaceae* (S Table 2). The taxa that revealed a significant effect of the cultivar are highlighted in Table 3 with "c" (c: 0.05-0.01; c*: 0.01-0.001; c**: < 0.001). A highly significant effect to the cultivar ($p < 0.001$) was only observed for *Streptomycetaceae*.

Clone library-based bacterial community analysis

Bacterial 16S rRNA genes were amplified from total rhizosphere DNA of the cultivars 'Baltica', 'Désirée' and 'Sibu', cloned and sequenced to taxonomically classify the abundant ribotypes in the rhizosphere of the three cultivars grown at two field sites. Analysis of the clones included only those with an average length of at least 400 nucleotides, resulting in a total of 151 ribosomal RNA gene sequences analyzed for the Roggenstein site as well as 160 sequences for the Oberviehhausen site. DOTUR analysis revealed that the rarefaction curves were still steeply growing at a distance of 5% suggesting that many more sequences would be needed to represent the diversity at this level. At a level of 30% the curves reached a plateau indicating that dominant phyla and classes were already detected. The majority of cloned 16S rRNA gene sequences in the six libraries had the highest sequence identity to sequences of uncultured bacteria. Overall, a total of 12 phyla and 17 classes were identified. The highest number of cloned 16S rRNA gene sequences belonged to the *Proteobacteria* (181/312). At both sites the highest numbers of proteobacterial 16S RNA gene sequences were assigned to the *Alphaproteobacteria* (99/181) (Table 5) followed by the *Betaproteobacteria* (55/181) followed by *Gammaproteobacteria* (24/181). Interestingly, there were more clones affiliated to the phylum *Acidobacteria* (40/312) than to the *Gammaproteobacteria*. Furthermore, sequences affiliated to the phyla *Bacteroidetes* (20/312), *Firmicutes* (19/312), *TM7* (16/312), *Actinobacteria* (14/312), *Verrucomicrobia* (8/312), *Gemmatimonadetes* (5/312), *Planctomycetes* (5/312),

Chloroflexi (2/312), *Cyanobacteria* (1/312) and *OP10* (1/312) contributed to the libraries. Clones affiliated to the phyla *Planctomycetes*, *Chloroflexi* and *Cyanobacteria* were only found in the rhizosphere of potato plants grown in Oberviehhausen. Sequences affiliated to *Rhizobiaceae*, *Sphingomonadaceae*, *Xanthomonadaceae*, *Acidobacteriaceae*, and TM7 were detected in the libraries of all three cultivars from both sites (Fig. 4). Other taxa such as *Bradyrhizobiaceae* were detected in all three libraries from the Roggenstein site but in none of the libraries from the cultivars grown in Oberviehhausen, while *Planctomycetaceae* were found in the Oberviehhausen libraries only.

The comparison of the libraries generated for both sites using the LibCompare tool of the RDP database with a confidence threshold of 95 % did not reveal significant differences except for the family *Oxalobacteraceae* ($p = 0.0009$) which was significantly more frequently detected in the clone library of the Roggenstein site.

Discussion

To date, several studies investigated how different factors such as the site, the plant species and the cultivar shape the composition of bacterial communities in the rhizosphere of potato plants (Lottmann *et al.*, 2000; Smalla *et al.*, 2001; Heuer *et al.*, 2002; Milling *et al.*, 2004; Rasche *et al.*, 2006; Becker *et al.*, 2008; van Overbeek and van Elsas, 2008). In most of these studies cultivation-independent 16S rRNA gene fingerprints such as DGGE or T-RFLP were used to compare the effect of different environmental factors on the bacterial community composition. The composition of bacterial communities in the rhizosphere of potato plants was shown to be dynamic and influenced by the plant growth development stage, the year and the site. Although several bands were characterized by cloning and sequencing (Lottmann *et al.*, 2000; Heuer *et al.*, 2002), the taxonomic composition of potato rhizobacteria and how their abundance was affected by the plant genotype or the site remained largely unknown. In this study the potential PhyloChip hybridization was exploited to obtain insights into the type of bacteria residing in the vicinity of potato roots and to identify those that respond to the cultivar and the site. As three independent field replicates per cultivar and site were analyzed rigorous statistical testing could be done. The high density 16S rRNA PhyloChips have not only a resolution of almost 10^4 taxa but also a range of detection that spans more than five orders of magnitude (Brodie *et*

al., 2006). However, in contrast to DGGE 16S rRNA gene fingerprints that are assumed to reflect the relative abundance of the dominant bacterial population, fluorescent signals cannot be taken as a measure of relative abundance. Instead the strength of the PhyloChip approach is that the fluorescence signals of each OTU detected can be compared among the different treatments. The total of 2,432 OTUs of the 8,743 resolvable taxa on the Phylochip that were detected in at least one of the replicates was in the same range as the number of taxa detected by DeAngelis *et al.* (2009) in the rhizosphere of the graminoid *Avena fatua* grown in three replicate soil microcosms. The data of the PhyloChip analysis suggest that a wide range of OTU differed in their relative abundance in the rhizosphere when the same cultivars were grown in two different soils. Thus, compared to DGGE fingerprints which were generated from the same DNA (Weinert *et al.*, 2009a) an improved resolution was achieved and taxonomic information on the major responders to site and the cultivar could be provided. Thus, the PhyloChip analysis could help to overcome the bias of molecular analysis methods described by Bent and Forney (2008) as the “tragedy of the uncommon” to some extent. However, less common ribotypes in the rhizosphere of potato plants (bacterial populations that are below 5 X 10^2 -10^3 cells per gram of soil) will be detected neither by DGGE nor by PhyloChip hybridization of 16S rRNA gene fragments amplified with *Bacteria* specific primers. The analysis of 16S rRNA gene fragments amplified from DNA extracted from rhizosphere bacteria uncovered an enormous diversity with more than 43 phyla detected. However, most of the OTU detected belonged to *Proteobacteria*, *Firmicutes*, *Actinobacteria*, *Bacteroidetes* and *Acidobacteria.* However, these numbers might be biased by the numbers of OTUs per phylum, class, order or family that can be theoretically detected. Furthermore, only ribotypes with high sequence similarity to 16S rRNA gene sequences which were present in the database at the time when this PhyloChip was developed will be detected. Therefore, we complemented the PhyloChip study with the cloning and sequencing of 16S rRNA gene fragments which were amplified from TC DNA. The latter approach offers the advantage of a resolution at the species level and enables the detection of novel ribotypes. Although the PCR conditions for the amplification of 16S rRNA gene fragments for PhyloChip hybridization slightly differed from those used for the amplification of 16S rRNA gene fragments for cloning and sequencing, all major phyla were detected by both approaches. However, the PhyloChip approach enabled us to detect a total of 43 phyla while only 12 were detected in the 16S rRNA

gene clone libraries. While *Acidobacteria* were the second most frequently observed phylum in the clone libraries, *Acidobacteria* were much less frequently detected by means of the PhyloChip. It is supposed that the full range of *Acidobacteria* diversity is not yet represented at the PhyloChip that was designed based on the sequences that were in the database 2005 (Brodie et al., 2006). OTUs belonging to the *Gammaproteobacteria* were the most frequently detected proteobacterial OTU on the PhyloChip. However, cloned 16S rRNA gene sequences affiliated to the *Gammaproteobacteria* contributed only 7.7 % of the sequences detected in the clone libraries. We suppose that due to a lower abundance of OTUs belonging to the *Gammaproteobacteria* these were less frequently represented in the clone libraries but their presence was still detected by the PhyloChip. *Epsilonproteobacteria* were not at all presented amongst the cloned 16S rRNA genes but still their presence was detected by means of the PhyloChip. None of the previously reported 16S rRNA gene sequences from potato rhizosphere was affiliated to *Epsilonproteobacteria* (Heuer *et al.,* 2002; Rasche *et al.,* 2006). *Alphaproteobacteria* represented the most abundant ribotypes in the rhizosphere of all three cultivars at both sites as the highest proportion of 16S rRNA gene sequences in all of the six libraries was affiliated to the *Alphaproteobacteria. Alphaproteobacteria* also comprised the most abundant ribotypes in the clone libraries generated from 16S rRNA genes amplified from TC DNA from grass, oilseed rape, and potato (McCaig *et al*., 1999; Kaiser *et al*., 2001; Heuer *et al*., 2002). While sequences with high similarity to *Bradyrhizobium* were obtained from all three cultivars grown in Roggenstein, these sequences were not detected amongst the libraries generated from the same cultivars grown in Oberviehhausen. In contrast, the study by Rasche *et al.* (2006) revealed a different taxonomic composition of 16S rRNA gene clones from the rhizosphere of genetically modified potato lines (derived from the cultivar ‘Désirée’) which were grown in Chernozem under greenhouse conditions with the majority of sequences affiliated to the divisions *Holophaga/Acidobacterium* (32-42 %), high-G+C Gram-positives (5-32 %), *Alphaproteobacteria* (4-27 %), *Betaproteobacteria* (3-18 %), and *Gammaproteobacteria* (3-11 %). Reasons for the contrasting results might be the different soil type but also the fact that in the study by Rasche *et al*. (2006) the rhizosphere was defined as soil which was brushed from the root after removing loosely adhering soil by shaking. Thus bacteria from soil more tightly adhering to the rhizoplane and to fine roots might have been absent in the rhizosphere community

DNA of the Rasche *et al.* (2006) study. However, in our study the relative proportion of phyla or classes detected by means of PhyloChips or by cloning and sequencing was remarkably similar for all cultivars grown at the two sites with rather different soil characteristics (Weinert *et al*., 2009a). Our findings also supported the observations by Vandenkoornhuyse *et al*. (2007) and DeAngelis *et al.* (2009) that bacterial phyla that are only rarely or not at all represented by isolates can consistently be detected in the rhizosphere. Thus phyla affiliated to the *Acidobacteria*, the *Verrucomicrobia* or *TM7* were detected at both sites from all cultivars (except *Verrucomicrobia* that were not detected in the clone libraries for Sibu). The role of these phyla in the rhizosphere remains to be explored but as shown by stabile isotope probing, e.g. *Acidobacteria* respond to root exudates (Vandenkoornhuyse *et al*., 2007). Highly significant effects of the site on the diversity in the rhizosphere were seen by PhyloChip analysis with few exceptions at different taxonomic levels. These observations confirm previous results by DGGE analysis of 16S rRNA gene fragments amplified from the same TC DNA (Weinert *et al.*, 2009a). However, the comparative analysis of the clone libraries from both sites did not reveal a significant effect of the site. Furthermore, the PhyloChip analysis also confirmed the recently reported conclusion from DGGE analysis that the effect of the cultivar is much less pronounced than the effect of the site. However, in contrast to DGGE the PhyloChip analysis enabled us to test at different taxonomic levels for statistically significant cultivar effects. The most striking findings from MANOVA testing were that the orders *Pseudomonadales* and *Actinomycetales* significantly responded to the cultivar. In particular *Streptomycetaceae*, *Micromonosporaceae* and *Pseudomonaceae* showed the strongest response to the cultivar. The *in vitro* antagonistic activities and biocontrol effects have been frequently studied for strains belonging to these families (Haas and Defago, 2005; Raaijmakers et. al., 2008; Berg and Smalla, 2009).However, also a relatively high proportion of OTUs belonging to the family *Enterobacteriaceae* displayed a significant response to the cultivar. OTUs affiliated to the *Enterobacteriaceae* were more abundant in the rhizosphere of the cultivar ‘Désirée’. The finding that OTUs of the *Enterobacteriaceae* contributed to the differences among cultivars is highly interesting, as many bacterial pathogens belong to this family (Stavrinides, 2009) and in particular members of the genera *Dickeya* and *Pectobacterium* (formerly Erwinia) are responsible for severe losses of tubers. It might be speculated that the importance of *Enterobacteriaceae* might become even

more pronounced at later plant growth stages. Indeed, in other studies, differentiating bands excised, cloned and sequenced from DGGE profiles of potato rhizosphere samples taken at EC90 were identified as *Enterobacteriaceae* species (Heuer *et al.*, 2002; Lottmann *et al.*, 1999). However, none of the OTUs belonging to the cultivar differentiating *Enterobacteriaceae* detected by PhyloChip analysis were detected in the clone libraries, most likely because the *Enterobacteriaceae* did not belong to the dominant ribotypes. The limited number of clones analyzed per cultivar and site might be the reason why a significant site effect was not detected when comparing the cloned 16S rRNA gene sequences by LibCompare, while both DGGE and PhyloChip analysis of PCR-amplified 16S rRNA genes revealed a significant influence of the site on the composition of the bacterial communities in the potato rhizosphere.

A much less complex 16S rRNA based taxonomic microarray was also used by Sanguin *et al.* (2006) to study the rhizosphere effect of maize on bacterial communities. The set of 170 probes on this array mainly targeted *Proteobacteria.* Probes targeting *Alphaproteobacteria*, *Deltaproteobacteria*, *Actinobacteria*, *Acidobacteria*, *Planctomycetes*, and *Bacteroidetes* yielded strong hybridization signals. Hybridization signals were observed for several *Alphaproteobacteria* probes, such as those targeting the three families *Rhizobiaceae*, *Bradyrhizobiaceae*, *Brucellaceae* and *Sphingomonas* spp. Also Sanguin *et al.* (2006) generated 16S rRNA gene libraries for rhizosphere and bulk soil to complement the microarray data. Interestingly, the proportion of clones belonging to the *Proteobacteria* was even higher in their study compared to our study (97% compared to 62% for the libraries from Roggenstein and 53% for Oberviehhausen). Although the majority of OTUs in the rhizosphere of all potato cultivars from both sites were affiliated to the *Proteobacteria,* the detection of OTUs with high sequence similarity to the candidate division TM7 or the *Acidobacteria* is quite remarkable. The potential role of these taxa in the rhizosphere is still a black box. The use of stable isotopes (^{13}C, ^{15}N) might help to better link photosynthates with bacterial responders and elucidate the role of ribotypes belonging to candidate divisions or taxa which are rarely represented in isolate collections (Wu *et al*., 2008; Rasche *et al.,* 2009).

Acknowledgement

This work was funded by grant 0313277B from the Bundesministerium für Bildung und Forschung. The authors would like to thank J. Dennert and F. X. Maidl (TU

München) for the perfect management of the experimental plots in Roggenstein and Oberviehhausen and are highly thankful to G. Wenzel (TU München) for providing the plant material of the transgenic lines. Ilse-Marie Jungkurth (Julius Kühn-Institut Braunschweig) is gratefully acknowledged for critically reading the manuscript.

References

Becker R, Behrendt U, Hommel B, Kropf S, Ulrich A. (2008). Effects of transgenic fructan-producing potatoes on the community structure of rhizosphere and phyllosphere bacteria. *FEMS Microbiol Ecol* **66:** 411-425.

Bent SJ, Forney LJ. (2008). The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. *ISME J* **2:** 689-695.

Berg G, Smalla K. (2009) Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* **68:** 1–13.

Berg G, Roskot N, Steidle A, Eberl L, Zock A, Smalla K. (2002). Plant-dependent genotypic and phenotypic diversity of antagonistic rhizobacteria isolated from different Verticillium host plants. *Appl Environ Microbiol* **68:** 3328–3338.

Berg G, Opelt K, Zachow C, Lottmann J, Götz M, Costa R, Smalla K. (2006). The rhizosphere effect on bacteria antagonistic towards the pathogenic fungus Verticillium differs depending on plant species and site. *FEMS Microbiol Ecol* **56:** 250-261.

Bremer C, Braker G, Matthies D, Reuter A, Engels C, Conrad R. (2007). Impact of plant functional group, plant species, and sampling time on the composition of nirK-type denitrifier communities in soil. *Appl Environ Microbiol* **73:** 6876-6884.

Briones AM, Okabe S, Umemiya Y, Ramsing N-B, Reichardt W, Okuyama H. (2002). Influence of different cultivars on populations of ammonia-oxidizing bacteria in the root environment of rice. *Appl Environ Microbiol* **68:** 3067-3075.

Brodie EL, DeSantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL, Hazen TC, Richardson PM, Herman DJ, Tokunaga TK, Wan JM, Firestone MK. (2006). Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl Environ Microbiol* **72:** 6288-6298.

Brodie EL, DeSantis TZ, Moberg Parker JP, Zubietta IX, Piceno YM, Andersen GL. (2007). Urban aerosols harbor diverse and dynamic bacterial populations. *Proc Natl Acad Sci* **104:** 299-304.

Costa R, Götz M, Mrotzek N, Lottmann J, Berg G, Smalla K. (2006). Effects of site and plant species on rhizosphere community structure as revealed by molecular analysis of microbial guilds. *FEMS Microbiol Ecol* **56:** 236-249.

Costa R, Gomes NCM, Krögerrecklenfort E, Opelt K, Berg G, Smalla K. (2007). Pseudomonas community structure and antagonistic potential in the rhizosphere:

insights gained by combining phylogenetic and functional gene-based analyses. *Environ Microbiol* **9:** 2260-2273.

Cruz-Martínez K, Suttle KB, Brodie EL, Power ME, Andersen GL, Banfield JF. (2009). Despite strong seasonal responses, soil microbial consortia are more resilient to long-term changes in rainfall than overlying grassland. *The ISME Journal* **3:** 738-744.

DeAngelis M, Brodie EL, DeSantis TZ, Andersen GL, Lindow SE, Firestone MK. (2009). Selective progressive response of soil microbial community to wild oat roots. *The ISME Journal* **3:**168-178.

Fisher MM, Triplett EW. (1999) Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to fresh water bacterial communities. *Appl Environ Microbiol* 65: 4630-4636.

Forney LJ, Zhou X, Brown CJ. (2004). Molecular microbial ecology: land of the one-eyed king. *Curr Opin Microbiol* **7:** 210-220.

Glimm E, Heuer H, Engelen B, Smalla K, Backhaus H. (1997). Statistical comparisons of community catabolic profiles. *J Microbiol Meth* **30:** 71-80.

Haas D, Défago G. (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nature Rev Microbiol.* **3:** 307-319.

Hack H, Gal H, Klemke T, Klose R, Meier U, Stauß R, Witzenberger A. (1993). Phänologische Entwicklungsstadien der Kartoffel (Solanum tuberosum L.). *Nachrichtenbl Dtsch Pflanzenschutzd* **45:** 11-19.

Heuer H, Kroppenstedt RM, Lottmann J, Berg G, Smalla K. (2002). Effects of T4 lysozyme release from transgenic potato roots on bacterial rhizosphere communities are negligible relative to natural factors. *Appl Environ Microbiol* 68: 1325-1335.

Heuer H, Kopmann C, Binh CTT, Top EM, Smalla K. (2009). Spreading antibiotic resistance through spread manure: characteristics of a novel plasmid type with low %G+C content. *Environ Microbiol* **11:** 937-949.

Kaiser O, Pühler A, Selbitschka W. (2001). Phylogenetic analysis of microbial diversity in the rhizoplane of oilseed rape (cv. Westar) employing cultivation-dependent and cultivation-independent approaches. *Microb Ecol* **42:** 136-149.

Kowalchuk GA, Buma DS, de Boer W, Klinkhamer PGL, van Veen JA. (2002). Effects of above-ground plant species composition and diversity on the diversity of soil-borne microorganisms. *Antonie van Leeuwenhoek* **81:** 509–520.

Kowalchuk GA., Drigo B, Yergeau E, van Veen JA. (2006). Assessing bacterial and fungal community structure in soil using ribosomal RNA and other structural gene markers. In: Nannipieri P, Smalla K (eds). Nucleic Acids and Proteins in Soil. Springer-Verlag: Berlin Heidelberg, Germany, pp 159-188.

Lemanceau P, Corberand T, Gardan L, Latour X, Laguerre G, Boeufgras J-M, Alabouvette C. (1995). Effect of two plant species, flax (Linum usitatissinum L.) and tomato (Lycopersicon esculentum Mill.), on the diversity of soilborne populations of fluorescent pseudomonads. *Appl Environ Microbiol* **61:** 1004-1012.

Lottmann J, Heuer H, Smalla K, Berg G. (1999). Influence of transgenic T4-lysozyme-producing potato plants on potentially beneficial plant-associated bacteria. *FEMS Microbiol Ecol* **29:** 365-377.

Lottmann J, Heuer H, de Vries J, Mahn A, Düring K, Wackernagel W, Smalla K, Berg G. (2000). Establishment of introduced antagonistic bacteria in the rhizosphere of transgenic potatoes and their effect on the bacterial community. *FEMS Microbiol Ecol* **33:** 41-49.

Marschner P, Yang C-H, Lieberei R, Crowley DE. (2001). Soil and plant specific effects on bacterial community composition in the rhizosphere. *Soil Biol Biochem* **33:** 1437-1445.

McCaig AE, Glover LA, Prosser JI. (1999). Molecular analysis of bacterial community structure and diversity in unimproved and improved upland grass pastures. *Appl Environ Microbiol* **65:** 1721-1730.

Milling A, Smalla K, Maidl FX, Schloter M, Munch JC. (2004). Effects of transgenic potatoes with an altered starch composition on the diversity of soil and rhizosphere bacteria and fungi. *Plant Soil* **266:** 23-39.

Muyzer G, Smalla K. (1998). Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek* **73:** 127-141.

Raajmakers JM, Paulitz TC, Steinberg C, Alabouvette C, Moenne-Loccoz Y. (2008). The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant Soil* **321:** 341-361.

Rasche F, Hödl V, Poll C, Kandeler E, Gerzabek MH, van Elsas JD, Sessitsch A. (2006). Rhizosphere bacteria affected by transgenic potatoes with antibacterial activities compared with the effects of soil, wild-type potatoes, vegetation stage and pathogen exposure. *FEMS Microbiol Ecol* **56:** 219-235.

Rasche F, Lueders T, Schloter M, Schaefer S, Buegger F, Gattinger A, Hood-Nowotny RC, Sessitsch A. (2009). DNA-based stable isotope probing enables the identification of active bacterial endophytes in potatoes. *New Phytologist* **181:** 802-807.

Sanguin H, Remenant B, Dechesne A, Thioulouse J, Vogel TM, Nesme X, Moenne-Loccoz Y, Grundmann GL. (2006). Potential of a 16S rRNA-based taxonomic microarray for analyzing the rhizosphere effects of maize on Agrobacterium spp. and bacterial communities. *Appl Environ Microbiol* **72:** 4302-4312.

Smalla K, Wieland G, Buchner A, Zock A, Parzy J, Kaiser S, Roskot N, Heuer H, Berg G. (2001). Bulk and rhizosphere bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Appl Environ Microbiol* **67:** 4742-4751.

Smalla K, Oros-Sichler M, Milling A, Holger H, Baumgarte S, Becker R, Neuber G, Kropf S, Ulrich A, Tebbe CC. (2007). Bacterial diversity of soils assessed by DGGE, T-RFLP and SSCP fingerprints of PCR-amplified 16S rRNA gene fragments: do different methods provide similar results? J Microbiol Meth **69:** 470-479.

Stavrinides J. (2009) Origin and evolution of phytopathogenic bacteria. In: Jackson RW (ed). Plant pathogenic bacteria – genomics and molecular biology. Caister Academic Press, Norfolk, UK, pp 1-35.

Vandenkoornhuyse P, Mahe S, Ineson P, Staddon P. (2007). Active root-inhabiting microbes identified by rapid incorporation of plant-derived carbon into RNA. *Proc Natl Acad Sci USA* **104:** 16970-16975

van Overbeek L, van Elsas JD. (2008). Effects of plant genotype and growth stage on the structure of bacterial communities associated with potato (*Solanum tuberosum* L.). FEMS Microbiol Ecol 64: 283-296.

Weinert N, Meincke R, Gottwald C, Heuer H, Schloter M, Berg G, Smalla K. (2009a) Rhizosphere microbial communities of genetically modified zeaxanthin-accumulating potato plants and their parent cultivar differed less than those of different potato cultivars. *Appl Environ Microbiol* **75:** 3859-3865.

Weinert N, Meincke R, Gottwald C, Radl V, Dong X, Schloter M, Berg G, Smalla K (2009b) Effects of genetically modified potatoes with increased zeaxanthin content on the abundance and diversity of rhizobacteria with in vitro antagonistic activity do not exceed natural variability among cultivars. *Plant and soil* DOI 10.1007/s11104-009-0024-z

Weisburg WG, Barns SM, Pelletier DA, Lane DJ. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173:** 697-703.

Wilson KH, Wilson WJ, Radosevich JL, DeSantis TZ, Viswanathan VS, KuczmarskiTA, Andersen GL. (2002). High-density microarray of small-subunit ribosomal DNA probes. *Appl Environ Microbiol* **68:** 2535-2541.

Wu WX, Liu W, Lu HH, Chen YX, Devare M, Thies J. (2008). Use of ^{13}C labeling to assess C-partitioning in transgenic and non-transgenic (parental) rice and their rhizosphere soil microbial communities. *FEMS Microbiol Ecol* **67:** 93-102.

Table 9 Number of OTUs detected in rhizosphere samples from different cultivar at two sites by phylochip

phylum	class	order	family	siteO			siteR			total
				Bal	Des	Sib	Bal	Des	Sib	
Proteobacteria	Alphaproteobacteria	*Bradyrhizobiales*		40±2.6	43±7.1	45±4	55±5.1	51±6	48±9.5	62
		Rhodobacterales		18±2.9	22±10.5	24±3.5	38±10.4	27±14.8	25±11.5	50
		Rhizobiales		44±7.9	47±14.4	52±5.3	65±3.6	62±6	57±9.1	76
		Sphingomonadales		36±4	38±13.2	43±4	54±2.9	51±5.6	48±7.1	66
		others		57±5.9	60±22.1	66±6.4	82±5.5	74±8.3	71±14.2	97
	Betaproteobacteria	*Burkholderiales*	*Comamonadaceae*	50±4.9	52±11	54±4	63±2.9	62±5	57±7	71
		Burkholderiales	*Burkholderiaceae*	9±2.1	10±3.5	11±4.2	17±0.6	15±2.5	13±3.5	19
			others	26±0.6	25±2.1	27±2.1	33±1	32±4.5	30±4.2	37
		Rhodocyclales	*Rhodocyclaceae*	11±3.2	13±5.3	15±2.6	21±2.5	17±4	16±6.6	24
		others		24±5.5	26±9.5	30±3.1	42±5.9	36±6.6	31±5.9	52
	Gammaproteobacteria	*Enterobacteriales*		49±15.6	66±8	66±5.5	72±5.3	71±6.4	53±24.3	87
		Pseudomonadales	*Pseudomonadaceae*	34±3.8	36±4	37±1.5	39±2.3	39±3.5	33±4.2	45
			Moraxellaceae	5±1.2	6±1.5	7±0.6	7±1	7±0.6	7±0.6	8
		Xanthomonadales	*Xanthomonadaceae*	11±1	13±5	14±2.3	23±2.6	19±4.9	17±6.6	29
		others		120±28.9	139±48.7	154±20	180±17	168±31.5	138±47.5	230
	others			120±14.2	125±25.1	133±11.1	148±3	136±14.2	132±25.1	171
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	14±0.6	15±3.1	15±1.2	18±2	18±5.2	16±3.2	25
			Micromonosporaceae	16±0.6	15±2.1	16±0.6	18±1.7	17±1.2	17±1	19
			Streptomycetaceae	15±0	15±0	15±0.6	17±1.2	15±0.6	16±1	18
			Micrococcaceae	18±1	18±1.7	19±0	20±0.6	19±1.2	20±0.6	21
			Pseudonocardiaceae	5±0.6	4±1.2	6±1.2	8±1.2	9±0	8±3.2	10
			others	88±9.9	88±24	101±7.2	114±5.7	103±6.5	97±23.6	128
		others		25±2.1	27±7	31±6.2	35±2.5	29±3.6	31±9	46
	BD2-10group			1±0	1±0.6	1±0.6	2±0	1±0.6	2±0.6	2
Firmicutes	*Bacilli*	*Bacillales*		91±19.2	89±32.6	104±4.4	117±6.4	105±13.1	96±23.8	147
		Lactobacillales		26±3.8	33±12.7	34±4.2	41±4.2	35±9.3	31±12.1	55
	Clostridia			113±29.6	112±50.1	131±9.2	162±14.3	151±22.7	132±32.9	196
	others			22±2	26±8	27±5.9	35±3.2	30±6.5	28±8.9	42
Acidobacteria	*Acidobacteria*	*Acidobacteriales*	*Acidobacteriaceae*	30±0.6	29±5.6	32±1.2	34±0.6	31±4	31±4.9	37
	others			27±2.3	25±8.5	29±4.2	31±0.6	25±6	28±7.2	38
Bacteroidetes	*Bacteroidetes*	*Bacteroidales*		18±5.5	24±8.2	24±3.5	25±3.8	28±7.9	21±5.8	43
	Flavobacteria	*Flavobacteriales*		31±3.5	33±6	35±5.5	42±4.4	40±5.6	32±8.1	58
	Sphingobacteria	*Sphingobacteriales*		40±5.9	40±15.1	46±7	52±7.2	49±6.5	43±13.1	70
	others			4±0.6	3±0.6	3±0	5±0.6	4±0	4±0.6	5
Spirochaetes	*Spirochaetes*	*Spirochaetales*		22±7	25±13.1	27±9.5	34±0.6	26±5.9	25±14.2	37
others				183±28.6	195±73.7	216±34.5	263±19.1	237±29.1	211±54	311
sum				1444±221.7	1541±461.3	1689±170.2	2015±138.2	1843±251.5	1664±410.1	2432

Note, an OTU was regarded as present in the sample when over 90% of its assigned probe pairs was positive (pf>=0.9).

Table 2: Numbers of OTUs detected specifically at high frequency for each site

phylum	class	order	siteO	SiteR
Proteobacteria	*Alphaproteobacteria*	*Rhizobiales*	0	12
		Bradyrhizobiales	0	8
		Sphingomonadales	0	7
		others	0	13
	Betaproteobacteria	*Burkholderiales*	0	15
		others	0	10
	Gammaproteobacteria	*Xanthomonadales*	0	6
		Alteromonadales	1	5
		others	4	11
	others		0	13
Actinobacteria	*Actinobacteria*		1	19
	BD2-10group	Unclassified	0	1
Firmicutes	*Bacilli*		3	9
	Clostridia	*Clostridiales*	0	15
	Mollicutes		0	6
others			8	44
sum			17	194

Note: specifically at high frequency: SiteO >=6/9 detected in site O while <=3/9 detected in Site R, vice versa for Site R

Table3: Summary of responding OTUs to treatments

phylum	Class	order	family	site	cultivar	site§cultivar	total
Proteobacteria	*Alphaproteobacteria*	*Bradyrhizobiales*		32(51.6%)	3(4.8%)	1(1.6%)	62
		Rhodobacterales		12(24%)	2(4%)	2(4%)	50
		Rhizobiales [c]		20(26.3%)	2(2.6%)	1(1.3%)	76
		Sphingomonadales		29(43.9%)	3(4.5%)	3(4.5%)	66
		others		22(22.7%)	1(1%)	0(0%)	97
	Betaproteobacteria	*Burkholderiales*	*Comamonadaceae*	39(54.9%)	3(4.2%)	3(4.2%)	71
			Burkholderiaceae	14(73.7%)	0(0%)	0(0%)	19
			others	14(37.8%)	1(2.7%)	1(2.7%)	37
		Rhodocyclales	*Rhodocyclaceae*	10(41.7%)	0(0%)	0(0%)	24
		others		13(25%)	5(9.6%)	2(3.8%)	52
	Gammaproteobacteria [c*]	*Enterobacteriales*		5(5.7%)	15(17.2%)	1(1.1%)	87
		Pseudomonadales [c]		15(28.3%)	21(39.6%)	7(13.2%)	53
		Xanthomonadales	*Xanthomonadaceae* [c]	12(41.4%)	1(3.4%)	1(3.4%)	29
		others		38(16.5%)	9(3.9%)	5(2.2%)	230
	others			43(25.1%)	9(5.3%)	2(1.2%)	171
Actinobacteria [c*]	*Actinobacteria* [c*]	*Actinomycetales* [c*]	*Microbacteriaceae*	11(44%)	2(8%)	2(8%)	25
			Micromonosporaceae [c]	3(15.8%)	10(52.6%)	1(5.3%)	19
			Streptomycetaceae [c**]	17(94.4%)	15(83.3%)	15(83.3%)	18
			Micrococcaceae	10(66.7%)	2(13.3%)	2(13.3%)	15
			Pseudonocardiaceae	6(60%)	3(30%)	2(20%)	10
			others	26(19.4%)	24(17.9%)	12(0,09)	134
		others		14(30.4%)	7(15.2%)	4(8.7%)	46
	BD2-10group			1(50.0%)	0(0.0%)	0(0)	2
Firmicutes [c]	*Bacilli*	*Bacillales*		53(36.1%)	17(11.6%)	15(10.2%)	147
		Lactobacillales		20(36.4%)	1(1.8%)	1(1.8%)	55
	Clostridia			45(0,23)	13(6.6%)	5(2.6%)	196
	others			12(28.6%)	5(11.9%)	4(9.5%)	42
Acidobacteria	*Acidobacteria*	*Acidobacteriales*	*Acidobacteriaceae*	13(35.1%)	2(5.4%)	1(2.7%)	37
	others			9(23.7%)	1(2.6%)	1(2.6%)	38
Bacteroidetes	*Bacteroidetes*	*Bacteroidales*		11(25.6%)	2(4.7%)	2(4.7%)	43
	Flavobacteria	*Flavobacteriales*		20(34.5%)	3(5.2%)	1(1.7%)	58
	Sphingobacteria	*Sphingobacteriales*		26(37.1%)	3(4.3%)	1(1.4%)	70
	others			1(20%)	1(20%)	0(0%)	5
Spirochaetes	*Spirochaetes*	*Spirochaetales*		8(21.6%)	2(5.4%)	2(5.4%)	37
others				68(21.9%)	19(6.1%)	9(2.9%)	311
sum				692(28.5%)	207(8.5%)	109(4.5%)	2432

Note: Responding OTU was identified by two way ANOVA with a unadjusted p value of 0.05 and summarized at different taxonomic levels cultivar effect was tested by MANOVA of the first five principal components c:0.01<p<0.05； c*:0.01<p<0.001; c**: p<0.001

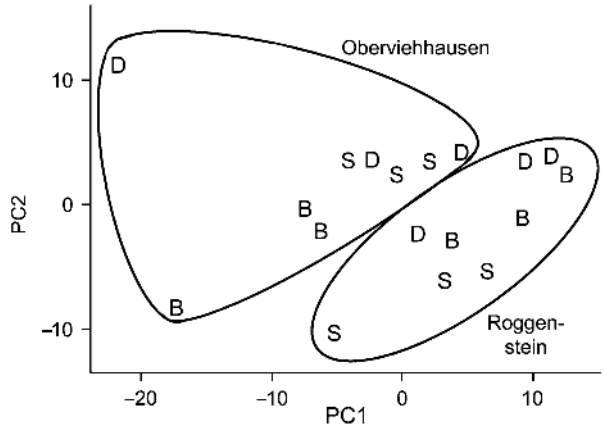

Fig. 1. PCA of the PhyloChip data from the potato rhizosphere samples of the cultivars ‘Baltica’ (B), ‘De´ sire´ e’ (D) and ‘Sibu’ (S) grown at the field sites in Roggenstein and Oberviehhausen. The first and second principal components represented 61% and 20% of the total variance.

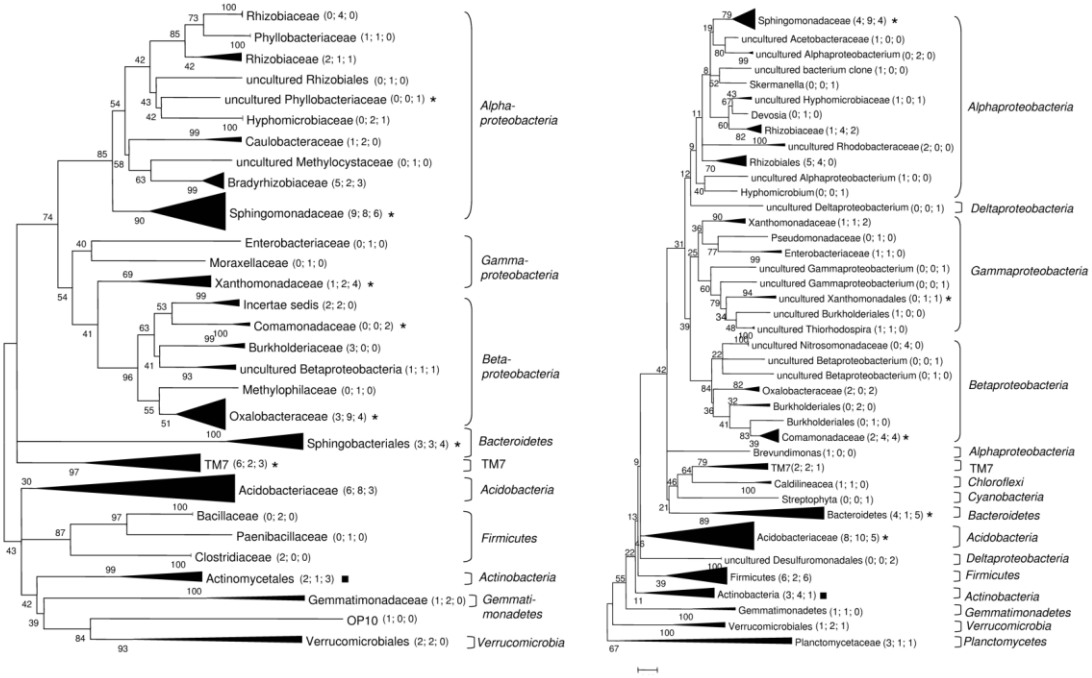

Fig. 2. Neighbour-joining trees of 16S rRNA genes amplified and cloned from rhizosphere samples of three potato cultivars grown at the field sites in Roggenstein (left) and Oberviehhausen (right). The numbers in parentheses refer to the number of sequences originating from the cultivars ‘Baltica’, ‘De´ sire´ e’ or ‘Sibu’, respectively. Asterisks indicate sequences that corresponded to OTUs detected by PhyloChip hybridization.